Volume 53 ▪ Number 1 ▪ March 2006

Editor-in-Chief ▪ FERENC KIEFER

FOUNDED IN 1951

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

WWW.AKADEMIAI.COM

AKADÉMIAI KIADÓ

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Acta Linguistica Hungarica publishes papers on general linguistics with particular emphasis on discussions of theoretical issues concerning Hungarian and other Finno-Ugric languages. Papers presenting empirical material must have strong theoretical implications. The scope of the journal is not restricted to the traditional research areas of linguistics (phonology, syntax and semantics, both synchronic and diachronic), it also covers other areas such as socio- and psycholinguistics, neurolinguistics, discourse analysis, the philosophy of language, language typology, and formal semantics.
The articles should be written in English, exceptionally, however, papers written in German and French are also accepted. The journal is published in yearly volumes of four issues.

■

Abstracted/indexed in
Bibliographie Linguistique/Linguistic Bibliography, International Bibliographies IBZ and IBR, Linguistics Abstracts, Linguistics and Language Behaviour Abstracts, MLA International Bibliography.

■

Editorial correspondence
ACTA LINGUISTICA HUNGARICA
Research Institute for Linguistics
P.O. Box 701/518, H-1399 Budapest, Hungary
Phone: (+36 1) 351 0413; Fax: (+36 1) 322 9297
E-mail: kiefer@nytud.hu; siptar@nytud.hu

■

Subscription price
for Volume 53 (2006) in 4 issues EUR 252 + VAT (for North America: USD 316) including online access and normal postage; airmail delivery EUR 20 (USD 25).

■

Please send your order to
AKADÉMIAI KIADÓ
Journals Business Centre
P.O. Box 245, H-1519 Budapest, Hungary
Phone: (+36 1) 464 8222; Fax: (+36 1) 464 8221
E-mail: journals@akkrt.hu
www.akademiai.com; www.akademiaikiado.hu

■

ISSN 1216-8076

ALing 53 (2006) 1

Printed in Hungary

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Volume 53, Number 1, March 2006

AKADÉMIAI KIADÓ
MEMBER OF WOLTERS KLUWER GROUP

Cover design: xfer grafikai műhely

ACTA LINGUISTICA HUNGARICA

VOLUME 53, NUMBER 1, 2006

CONTENTS

Acta Linguistica Hungarica, Vol. 53 (1), pp. 1–39 (2006)
DOI: 10.1556/ALing.53.2006.1.1

CHANGES IN ARGUMENT STRUCTURE IN THE COURSE OF DERIVATION IN HUNGARIAN*

GÁBOR ALBERTI

Department of Linguistics
University of Pécs
Ifjúság útja 6.
H–7624 Pécs
Hungary
albi@btk.pte.hu

Abstract: For all types of derivation characterised as productive by Kiefer (2000), the original version of Model Tau (Alberti 1997), dealing only with verbal derivation coming with no category change, can be extended to the entire spectrum of derivations; moreover, it can be extended in a straightforward way: the single novel factor is the **central case frame** peculiar to particular word categories. For instance, if the predicator is a noun, what corresponds to the case frame ⟨Nom, Acc⟩ in the sphere of verbs, is the case frame ⟨Nom, Poss⟩; this mapping is immediately observable in the case of *-ÓjA* (Laczkó 2000b), a suffix forming nouns in an argument-structure retaining way (*elcsábít* 'seduce' → *elcsábítója* '(someone's) seducer'). The case frame characteristic of the output word category supplies an upper limit, within which the actual realization can belong to five types that precisely coincide with the five basic types of category-preserving verbal (and participial) derivation discussed in Alberti (*ibid.*). How can these five basic types be derived? The crucial factor of each argument-structure transition is "advancement" of an argument (parallel with the "degradation" of another argument) in a sense that can be precisely defined in Model Tau.

Keywords: word formation, thematic theories, argument hierarchy, Model Tau

In the nineties I worked out a universal theory on potential case-frame manifestations of verbal argument structures and systematizing their

* Writing this paper has been sponsored by OTKA (No. T38386). Special thanks are due to the following colleagues for their valuable comments on earlier versions of this paper: Katalin É. Kiss, Emese Bérces, Magdolna Ohnmacht.

transitions by the name **Model Tau**, chiefly illustrated by Hungarian data and analyses (see e.g., Alberti 1988b,a, 1992–1993, 1994, 1997, 1998b), in which the holistic manner of Zsilka's thinking and his observations I judge to be the most valuable and opportune (e.g., Zsilka 1966, 1982) have been reconciled with the decisive elements of **thematic theories** (e.g., Komlósy 1992);[1] joining the trend according to which—instead of attempting to fix a particular list of thematic roles—morphosyntactically relevant lexical-semantic information is to be captured by means of more abstract **argument hierarchies** (e.g., Bresnan–Kanerva 1989; Grimshaw 1990; Dowty 1991).

This version of Model Tau, thus, contains statements—primarily constraints—on verbal derivations coming with no category change, sorting them out into five types of transition. This project aims at extending the system to all kinds of derivation claimed by the authors of the volume on (Hungarian) morphology of the *Strukturális magyar nyelvtan* series (Kiefer 2000) to be (at least semi-) productive. What is primarily claimed is that in this "extended Model Tau", into which cases of word-formation with non-verbal output or input have been integrated, a single novel factor should be allowed for in the course of describing argument structure changes—while retaining the hypothesis on the five sorts of transition types; nothing else but what is straightforward and inevitable: the case-frame possibilities characteristic of the category of output words.

Let us consider the content of the paper. After sketching and illustrating Model Tau (in section **1**), I expound the crucial principles that the system of Hungarian argument structure transitions shown in the Appendix relies on (section **2**), then distinct sections will be devoted to reviewing derivations **retaining**, **expanding** or **reducing** argument structure (sections **3**–**5**).

1. Absolute chain of influence—relative case frames

One of the crucial ideas that Model Tau relies on is that a given argument frame of a given verb form is to be characterized on the basis of a family of related argument structure versions to be assembled in a specific way, instead of the often uncertain methods based on lists of thematic roles.

[1] I would like to recall the prosperous and far-reaching family of thematic theories by mentioning the following seminal works: Gruber (1965); Fillmore (1968); Jackendoff (1987); Rappaport–Levin (1988); Parsons (1995); Williams (1995).

There are, for instance, at least five participants in the conceptual frame of the verb form *elásat* 'make sy dig sg into the ground', shown in (1d) below: in addition to the explicitly appearing privateer giving orders and the treasures getting in the depths of the ground, in certain argument structure versions (ASV) of *ás* 'dig' further participants can also be mentioned, such as the pirates carrying over the work of digging in fact, the pieces of soil moved (and, hence, directly affected) by them, and the holes accommodating the treasures. With regard to traditions, I have attempted to associate these roles with thematic-role labels, but in this way a potential characterization is such that the privateer, **enforcing** his will on others, and the pirates, who are **carrying out** some job, should both be qualified as **Agents**, and the pieces of soil, which are **moving**, the holes **appearing**, and the treasure, **placed** in the ground, are all regarded as **Patients**. As separation is a preferred aim, the privateer can be characterized as a **Causer**, and the holes can be understood as the **Goal** of a change of state pertaining to the ground of an area. I do not enter into details, I have only intended to point to the above mentioned uncertain nature of selecting thematic roles. What has been proposed in Model Tau as a relevant lexical-semantic characterization, is a **polarized chain of influence**, demonstrated in (1e) below; whose innovation lies in the fact that a given role in a given ASV belonging to a given verb is essentially to be calculated on the basis of a linearly ordered list of related ASVs (illustrated in (1a–d) below)). What can be said about this linear ordering here is that an appropriate list should consist of ASVs with richer and richer meaning content; the construction of such lists requires precise and quite intricate procedures as well as the algorithmic construction of polarized chains of influence does (their presentation would go beyond the scope of this paper, but see Alberti 1997).

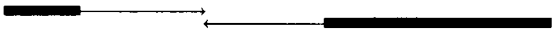
(1) (a) A kalózok (a szigeten) ásták a földet. ⟨Agent, Patient⟩
the pirate-pl (the island-sup) dig-past-3pl the soil-acc
'The pirates were digging the ground (on the island).'

(b) A kalózok gödröket ástak. ⟨Agent, Goal⟩
the pirate-pl hole-pl-acc dig-past-3pl
'The pirates were digging holes.'

(c) A kalózok elásták a rabolt kincseket. ⟨Agent, Patient_2⟩
the pirate-pl away-dig-past-3pl the stolen treasure-pl-acc
'The pirates buried the stolen treasures.'

(d) A kalózvezér *elásatta* a rabolt kincseket. ⟨Causer, Patient⟩
the privateer away-dig-cause-past-3pl the stolen treasure-pl-acc
'The privateer had the stolen treasures buried.'

(e) The chain of influence belonging to *elásat* with strong/weak Agent-like and Patient-like roles (on the basis of (1a–d)):
Causer → Agent → Patient → Goal → Patient_2



The initial step of the algorithm mentioned above is to mark the ASV illustrated in (1a) above as the **primitive core** of the chain of influence to be constructed; the argument occupying the subject/object position in this ASV is qualified as a **weak agent/weak patient** in our approach, respectively, indicated by thin arrow-like underline above in (1e). ASV (1b) provides a new argument in its object position: this will be the first **strong patient** (thick arrow-like underline). The new object in ASV (1c) is to be qualified as the successive strong patient in the chain of influence. ASV (1d) provides a new argument in its subject slot. In a situation like this the left edge of the chain of influence is to be enriched—with a **strong agent**.

The two fundamental (falsifiable) universal predictions of Model Tau exert **restrictions** on the subject and the object of certain transitive members of ASV families—as "relative" ASV selections: the chain of influence in the background and the qualification of certain arguments as strong agents or patients are to be "respected", i.e., the **absolute polarized chain of influence** is to be complied with. An object, thus, is **not** allowed to precede the subject serving as its co-argument in the same ASV according to the chain of influence in the background; further, a strong patient will **not** appear as the subject of a transitive member of the ASV family, and a strong agent will **not** serve as an object. These two fundamental constraints say nothing about arguments with a weak polarity (weak agents/weak patients); that is, Model Tau offers no categorical statements concerning arguments belonging to a primitive core (in the sense defined above). The reason lies in the observation that arguments in the primitive core show major (less predictable) variability in respect of realization as subjects or objects.[2]

[2] In the family partly demonstrated in (1), thus, there is no ASV (subcategorized by an arbitrarily affixed variant of the verb stem *ás* 'dig') in which, for instance, the argument that the role of holes is associated with would appear as the subject with the privateer's argument as the corresponding object (2a). Passive ASVs,

(2) Constraints on relative transitive ASVs belonging to an absolute chain of influence:

a. * Object (+) → ... → Subject (−)

b.1 * Object (+) ▬▬▬▬→

b.2 * Subject (−) ←▬▬▬▬

In the background of the falsifiable synchronic statements there is a psycholinguistic hypothesis to be verified statistically, which is completed with a typological conjecture, a hypothesis on language acquisition and another one on the evolution of human language, providing a coherent model.

According to the first hypothesis, an absolute chain of influence belonging to a particular verb stem is a true reflection of the flow of influence "in reality," at least in the following sense: the chain of influence in the grammar will coincide with the real flow of influence if this latter flow can be clearly appraised (e.g., in the realm of physical powers) while, otherwise, the chain of influence should be regarded as the outcome of a predominantly accidental process of grammaticalization. The minimal pair *please/like* can serve as an excellent illustration of the latter case (e.g., *Peter likes Mary* versus *Mary pleases Peter*): *like* has been grammaticalized in a way that the thematic frame ⟨Experiencer, Stimulus⟩ is mapped onto a case frame ⟨Nom, Acc⟩, whereas this assignment is in the reverse order in the case of *please*, yielding a case frame ⟨Acc, Nom⟩, despite the fact that it is not easy to indicate any difference to be considered relevant between the two meanings. In the relation between an Experiencer and a Stimulus there is no (inevitable) real physical impact. An Experiencer, on the one hand, can be taken as a participant whose feelings manifest themselves in the form of activities, but the role of a Stimulus can also be regarded in a way that triggering feelings is a result of different activities. In the case of *digging*, however, the flow of impact is essentially clear: digging people move pieces of soil, and not the other

unobjectionable in lots of languages, are not to be regarded as a violation of the prediction in question, as they are intransitive variants with the agentive argument in an oblique case (e.g., 'Holes were dug *by the pirates.*'). There is no transitive ASV, either, in which the holes would be selected to occupy a subject position with the argument of treasures as the corresponding object (2b.2). As for the possibility of a variant illustrated by the sentence ??*A kalózvezér már órák óta ásatja az embereit* 'The privateer has been making his people dig for hours.' (cf. (2b.1)), nothing is claimed about it, as the role of the people digging is only weakly agentive; which is in perfect harmony with the uncertain judgment concerning the grammaticality of the sentence just mentioned ("??").

way round; and holes are the result of the dislocation of these pieces of soil, so holes are "later" both in a logical and in a chronological sense. As for the relation between holes and treasures, this relation may be taken as an artificial impact, grammaticalized according to our human aspect. As for the side of causers, the privateer's influence is abstract, of course, but it is deeply rooted in the "secondary reality" of human group hierarchies.

The above mentioned hypothesis in the area of language typology says that there are two **central** argument functions in every natural language: a **negative** relative role (−) of agentive polarity, and a patient-like **positive** relative role (+). In our approach a **transitive** ASV is defined as their simultaneous occurrence, which is hypothesized to serve the purpose of revealing a fragment of the polarized chain of influence in the background. The constraint formulated in (2a), hence, can straightforwardly be understood as a consequence of this requirement demanding compatible representation. An **intransitive** ASV is regarded as such in which the same participant plays both the causer's negative relative role and the causee's positive relative role (−/+). What is highlighted in the case of a sentence like *The pirates have been digging for hours*, is the pirates' getting tired rather than the fact that pieces of soil have been moved or, probably, holes have been created. In the **accusative** type of languages (which English and Hungarian belong to), the negative central function and the positive central function are to be indicated by the unmarked Nominative and the morphologically (and/or syntactically) marked Accusative case, respectively, whilst the single argument of intransitive ASVs is in the Nominative case. This latter statement also holds true of the **ergative** type of languages, but in this type it is the positive central function that will remain case-morphologically unmarked (Absolutive case), whereas the negative central function will get an Ergative case. By introducing these two sorts of central polarity (Alberti 1997), not only these two language types can be covered but the entire rich realm of intermediate/mixed types reviewed by Komlósy (1982).

The hypothesis on language acquisition mentioned in an earlier paragraph predicts a straightforward connection between the absolute polarized chain of influence (in the background) and the relative ASV realizations ("on the surface"): an infant who is acquiring a language—parallel with learning what kinds of physical and psychic influences work in the world around him/her—will construct the polarized chains of influence corresponding to verbs on the basis of ASVs heard, and then (s)he will (be able to) construct new ASVs never heard by him/her. In the knowl-

edge of the data in (1) and the general constraints in (2), for instance, an infant is able to construct an ASV like the one shown in (3a) below, which is perfect indeed. Nevertheless, we should be careful about these positive predictions; it is better to speak about tendencies, which may be more or less reliable partly depending on different language-specific factors. This precaution concerns the possibility of intransitive ASVs as well (e.g., ??*Mostanában az a főnök heppje, hogy napokon át ásat* 'Nowadays the boss is hipped on making [people] dig for days.').

(3) (a) A kalózvezér gödröket ásatott. ⟨Agent, Goal⟩
the privateer hole-pl-acc dig-cause-past-3pl
'The privateer had holes dug.'

(b) ??Milan Bartoš berúgta a cseh válogatottat a torna
MB in-kick-past-3sg the Czech team-acc the competition
döntőjébe.
finals-poss3sg-ill
'The Czech team reached the finals due to a goal scored by MB.'

(c) ??Bocsi, édeském, most leteszlek/kinyomlak
sorry, sweet-poss-1sg, now down-set-s1sg-o2sg/out-press-s1sg-o2sg...
(mert úgy hallom, érkezik a férjem).
'Sorry, my dear, now I finish speaking to you [by hanging up the receiver/pressing the appropriate button of the mobile phone] because I hear my husband coming.'

The hypothesis on language evolution, also mentioned above, is a similar basic element of the philosophy that Model Tau relies on: it is assumed to be an important part of human adjustment to the current world that we strive for expressing connections which are "far away" in the physical reality but **relevant** to us by means of a single verb (and argument structure) , by connecting "distant" roles in this way. The privateer in (1d), for instance, does not touch the soil; it is even possible that he does not touch the treasures either. We can say, nevertheless, that the digging activity in the situation demonstrated in (1d) serves his purpose of putting the treasures in holes. The diachronic process can be detected in synchrony in the form of funny bloomers or unguarded wordings, like those in (3b–c) above. The famous Czech forward, for instance, kicked a ball, immediately, and in this way he could 'kick a goal'; which is of a crucial relevance in the course of a single match as well as in the course of a series of matches, whose finals can be reached for the whole Czech national team by means of such instances of scoring goals. Bartoš, hence, managed to exert a relevant impact on his team by kicking, immediately, a

ball. It is, then, a problem which can be modeled by stochastic processes from this point (Alberti 1988b,a) whether a new ASV will be naturalized in the language of a social layer or in a technical jargon, and then in the standard language, or this potential process will abort at birth. Similar thoughts can be formulated in connection with the example in (3c); what is the crucial point here is that the lady telephoning exerts some relevant influence on her interlocutor by finishing their conversation through immediately exerting influence on an object.

In what follows, methods of Model Tau are illustrated by two other ASV families, in (4) and (5) below. It is worth noticing that the same four thematic roles are mapped onto different case frames; we can conclude, hence, that traditional thematic representations cannot optimally capture the semantic character relevant from morphosyntactic points of view. The crucial difference lies in allowing for what has been called the **primitive** ASV, a proposal peculiar to Model Tau in this form.

(4) (a) Betört egy ablak. ⟨Patient⟩
break-past-3sg a window
'A window broke.'

(b) Ez a kalapács még egy vastag ablakot is betörne. ⟨Instr., Patient⟩
this the hammer even a thich window-acc even break-cond-3sg
'This hammer could break even a thick window.'

(c) Péter betörte az ablakot (egy kalapáccsal). ⟨Agent, Patient, (Instr.)⟩
Peter break-past-3sg the window-acc (a hammer-inst)
'Peter broke the window (with a hammer).'

(d) Mari *betörette* az ablakot Péterrel / egy kalapáccsal.
⟨Causer, Patient (Ag/In)⟩
Mary break-cause-past-3sg the window-acc Peter-inst / a hammer-inst
'Mary made Peter break the window.'/'Mary had the window broken with a hammer.'

(e) The polarized chain of influence belonging to *betöret* (on the basis of (4a–d)):
Causer → Agent → Instr → Patient

The intransitive ASV in (4a) above will serve as the primitive core of the ASV family on the basis of which the chain of influence belonging to the ASV in (4d) can be calculated. The Instrument (or Natural Force) which appears in variant (4b), thus, provides a new subject, to be identified

as a strong agent (according to the algorithm of calculation mentioned above). The Agent and the Causer in variants (4c–d) above provide further strong agents, resulting in the polarized chain of influence shown in (4e). In the ASV family demonstrated in (5) below, however, the primitive ASV in (5a) is transitive and, hence, the argument belonging to the prickly object is to be qualified as a weak agent. As a result of the constraint in (2b.1) above, thus, it is predicted that the entire ASV family whose members are mentioned in (4) will not contain an ASV that would correspond to the one shown in (5b) (cca. **Péter odatörte a kalapácsot az ablakhoz* [Peter towards-break-past-3sg the hammer-acc the window-all], intended meaning: 'Peter knocked the hammer against the window, and the hammer broke.').

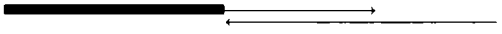
(5) (a) Egy szög megszúrta a kezemet. ⟨Patient, Goal⟩
a nail perf-prick-past-3sg the hand-poss3sg-acc
'A nail pricked my hand.'

(b) Péter beleszúrt egy szöget az abroncsba. ⟨Agent, Patient, Goal⟩
Peter into-prick-past-3sg a nail-acc the tyre-ill
'Peter pushed a nail into the tyre.'

(c) Péter *kiszúrta* az abroncsot egy szöggel. ⟨Agent, Pat/Ins, Goal/Pat⟩
Peter out-prick-past-3sg the tyre-acc a nail-inst
'Peter punctured the tyre with a nail.'

(d) Mari *kiszúratta* az abroncsot. ⟨Causer, Goal/Pat⟩
Mary out-prick-cause-past-3sg the tyre-acc
'Mary had tyre punctured.'

(e) The polarized chain of influence belonging to *kiszúrat* (on the basis of (5a–d)):
Causer → Agent → Instrument → Patient



At this point it is worth raising the dividing dilemma of thematic theories (e.g., Dowty 1991, section 5): whether the identifiable arguments of related ASVs can be associated with distinct thematic roles. Can it be said, for instance, that the prickly thing in ASV (5a) is a Natural Force, as it functions according to its natural character (like the wind blows), whereas in ASV (5b) it appears as a Patient, as it is moved towards a goal, and in ASV (5c) it is the Instrument of an Agent's exerting (fatal) influence on an earlier Goal, which qualifies as a Patient in this version in the spirit of this approach? Followers of this approach, thus, think

that participants to be identified play slightly modifying roles from ASV to ASV, explaining the abundant supply of ASVs that can be observed from language to language. This approach suffers from a cruel paradox, however: how can we consider the roles associated with different ASVs to be identical/identifiable? I prefer the view of the opposite party (Alberti 1997), which I judge to be traceable back to the original approach by Fillmore (1968) relying on deep cases: thematic labels of identical/identifiable roles should be constant, which does not exclude at all that the same thematic frame could be mapped onto different case frames. The problem with this approach is that in the course of an isolated examination of particular ASVs the traditional definition-like formulations (Komlósy 1992) present a basis for associating the "identical" roles with different thematic labels, as has been discussed above.

Model Tau, due to its "dynamic" approach, can essentially avoid the dividing dilemma discussed above, because in this theory arguments are not attempted to be characterized on the basis of "what is in the real world" but what has been grammaticalized in the language. An entire ASV family supplies the **absolute** information in the background on grammaticalized chains of influence and poles according to agentivity (where verb stems are fixed in a family but no rigid formal restrictions on affixation are applied), whilst particular ASVs, due to their peculiar case frames, can express slight **relative** "shift of emphasis": which are the participants whose connection the speaker considers to indicate the impact of influence relevant in the given situation, including reflexivity exhibiting intransitive ASVs.

2. Word formation and ASV transition coming with it: advancing and degrading arguments

My starting points in questions concerning word formation are the observations and definitions serving as commom denominators in the "structuralist" morphology edited by Kiefer (2000) and written by many authors.[3] My aim is not to rebuff or to correct their statements but to show that the original Model Tau can be extended in an optimally straightfor-

[3] The descriptive / systematizing chapters of the volume Kiefer (2000) are of principal importance to the topic of the present paper: Kenesei (2000); Kiefer–Ladányi (2000a,b); Laczkó (2000b,c); Komlósy (2000).

ward manner to all the cases qualified by the volume as (at least **semi-**) **productive** derivations.

The first question to face obviously concerns the nature of word formation. A sufficient criterion is **change of category**, whose crucial element can be recognized in the fact that the accommodating or refusing syntactic and/or morphological environments are not the same in the input phase and in the output phase of the given derivation. The type of derivation coming with **no category change** can be captured by modifying this criterion so that what is referred to will be the morphosyntactic environments that the word can (or cannot) **accommodate**, that is, practically case frames. The derivational connection between successive members of the ASVs of the families shown in (1), (4) and (5) above belongs to this latter type of word formation, as Model Tau in its original form (Alberti 1997) had been intended to account for derivational relations within the category of verbs. Word formation may come with (explicit) derivational affixation, which is an additional factor that sufficiently proves that derivation has taken place but is not necessarily to be regarded as an obligatory component of derivation: I agree (e.g., Alberti 1997) with the introduction of the concept of **conversion** (Kiefer–Ladányi 2000a,b; Laczkó 2000b), defined as the sort of derivation with no (explicit) morphological change of the input word form. In the least easily evaluable cases even changes in case frame should be dispensed with. In the case of the transition shown in (1a–b), for instance, the fact of derivation can be detected by indicating that the argument structure has been changed as certain argument slots are under the influence of a modified selectional restriction (Komlósy 1992) (the role of the "dug soil" can be occupied by solid things whereas the participant corresponding to the holes is "made of air").

As shown by the series of examples in (5) above, due to the "functional" definition of derivation (enriched with elements of a logical nature) peculiar to Model Tau, what can be accepted as a morphological modification serving as a formal concomitant to derivation is not only adjoining morphemes to a relative stem but substituting an affix for another one. That is, we can regard versions of a word which cannot be related by some productive morphological rule but are due to more or less accidental diachronic processes (Alberti 1997, 149) as standing in an input–output derivational relation. What is at issue is essentially the **blocking effect** discussed by Komlósy (1992), according to which, an otherwise productive morpheme-adjoining operation will not function where

both an appropriate input word form and an appropriate output word form exist, in the form of words coming from earlier states of the language (e.g., the causative version of *fő* 'be on the boil' is not *fövet* (cf. *lő* 'shoot' – *lövet*), but *főz* 'cook'; it is just the irregular variant *főz* that prevents the theoretically potential regular variant *fövet* from being used).

Model Tau has been intended to provide a descriptively and explanatorily adequate theory (and system) of possible transitions of **central frames** (Alberti 1997); thus, in the case of word formation, the question is as follows: which argument **was** and which argument **will be** in the nominative case, in the accusative case or in some oblique case. In the extended version of Model Tau, in which derivations coming with category change are also allowed for, this question may be changed as follows: what kinds of central case frames can be associated with different categories of words? Let us start, then, with reviewing this question, by casting the first glances at the chart in the Appendix, serving as the central topic of this paper.

As was established above, in the Hungarian language, which belongs to the accusative family of languages, the transitive case frame is such that the argument playing the causer's role, i.e., the one with the **negative** central role, will be associated with Nominative case, whereas the "causee", i.e., the argument with the **positive** central role, will get the Accusative; see (6) below. An intransitive case frame consists of a single Nominative, which we regard as associated with **both** central roles in a situation like this (6b). **Participles** essentially also use the central case frame ⟨NOM, ACC⟩, with the following differences: their subject-like argument typically gets no phonetic form ((6c): "∅"), although this possibility is not excluded in a few marginal or archaic constructions (6d); and finally in a construction with an infinitive in its center the subject can appear in a possessor-like way (6e). In the systematizing chart in the Appendix, derivations with a verbal or a participial output have been placed in the same column, with regard to the essential similarity of the central case frames. The reason of this similarity can be found in the fact that participles behave as transitional categories in a way that their output nominal nature is expected by their "accommodating environtment" in sentences, whereas environments that they can potentially accommodate, i.e., their argument strucutres, show their input (verbal) nature.

(6) (a) Mari megtelefonálta a hírt. ⟨NOM, ACC⟩ = ⟨–, +⟩
Mary perf-phone-past-3sg the news-acc
'Mary telephoned somebody about the news.'

(b) Mari telefonálgat. ⟨NOM⟩ = ⟨−/+⟩
Mary phone-dim-3sg
'Mary is telephoning [aimlessly].'

(c) [látva a helyzetet] / egy [sokakat piszkáló] fiú /
[see-adv the situation-acc] / a [many-pl-acc annoy-presprt] boy /
(Mit akarsz?) [Megtelefonálni a hírt Péternek.] ⟨∅, ACC⟩
(what-acc want-2sg) [perf-phone-inf the news-acc Peter-dat]
'[seeing the situation]' / 'a boy [annoying a lot of people]' / '(What do you want?) [To telephone Peter about the news.]'

(d) ?egy sokakat piszkáló, de [maga is sokak által piszkált]
a many-pl-acc annoy-prespar, but [self too many-pl by annoy-pastprt]
fiú / [bika rugaszkodván] / a [helyzet kínálta] lehetőség
boy / [bull push-off-adv] / the [situation offer-pastprt-poss3sg] facility
⟨NOM, ∅⟩
'a boy who annoys many other people but is [also annoyed by many people]'
'[As the bull pushed off,...]'/'the facility [offered by the situation]'

(e) (Tudod, mi volt a legnagyobb hiba?)
know-2sg what be-őast-3sg the greatest mistake
Marinak megtelefonálnia azt a hírt. ⟨POSS, ACC⟩
Mary-dat telephone-inf-3sg that-acc the news-acc
'(What was the greatest mistake?) For Mary to phone about that news.'

(7) (a) ?egy másokat részegséggel vádoló, közben [maga is] részeg /
an other-pl-acc drunkenness-inst accuse-presprt, whilst self too drunken /
Mari részeg. / Mari büszke a lányára. ⟨NOM⟩
Mary drunken / Mary proud the daughter-poss3sg-sub
'one who accuses others of drunkenness whilst he himself is also drunken' /
'Mary is drunken.' / 'Mary is proud of her daughter.'

(b) egy [részeg] tengerész / [Részegen] táncoltunk. ⟨∅⟩
a [drunken] sailor / [drunken-adv] dance-past-1pl
'a drunken sailor' / 'We danced in a state of drunkenness.'

(c) Ez sértés / a törvények durva megsértése! ⟨NOM, POSS⟩
this violation / the law-pl crude violation-poss3sg
'This is an offence / a crude violation of laws.'

(d) Nem tűröm a sértéseket / a törvények durva megsértését. ⟨∅, POSS⟩
not bar-1sg the offence-pl-acc / the law-pl crude violation-poss3sg-acc
'I will not have the offences / the crude violation of laws.'

The examples in (7) above are intended to illustrate the central case frame of ASVs associated with nominal predicators. It is excluded that

an argument belonging to a **nominal** predicator would be marked with Accusative. A predicator of the category adjective, also depending on its function, has a single central case, which is Nominative (7a); moreover, this argument position is typically associated with no phonetic form (7b). The latter situation is characteristic of the case frame of an adverbial predicator (7b), that is why the adjectival and adverbial derivational morphemes share a column in the Appendix. As for a predicator of the category noun, it will turn out that it is worth attributing to it a central frame in which the nominative case marking is associated with the negative central role (7c), or does not appear explicitly (7d), whereas the marker of the positive central role is the possessor function (7c–d). I give here the following observation as the first argument in favor of this approach: as a result of nominalization, the object of the input verb turns into a possessor: *megsérti a törvényt* → *a törvény(-nek a) megsértése* [perf-violate-3sg the law-acc → the law(-dat the) perf-violate-noun-poss3sg] 'violate the law' → 'violation of the law' (Alberti 1995).

As for the output case frame of a derivation, the case frame characteristic of the output word category only supplies an upper limitation, within which the actual realization can still belong to **five** types; and I claim that these five possibilities precisely coincide with the five basic types of category-preserving verbal (and participial) derivation discussed in the original 1997 framework of Model Tau (Alberti 1997). How can these magical five basic types be derived? The crucial factor of each ASV transition is claimed to be nothing else but the **advancement** of an argument, that is, the increase of its relative role. As will be gone over in (8) below (see the transition schemes listed in the left column), what can be regarded as an "advancement" in the case of a non-central argument is obtaining a central role (which is in accordance with its polarity) (8a–b), while a central participant can obtain an even higher position by acquiring both central roles (8c–d). In the case of a transitive input, which is to be regarded as the basic case, the advancement of an argument in the above discussed sense will necessarily come with the **degradation** of a central input argument. It can be checked that if we declare a principle requiring the possible "least change" in the course of a transition, according to which a single ASV transition can result in some change in the central role of **at most two arguments** (i.e., no argument besides the argument "to be advanced", or one argument if and only if this change is inevitable because of the universal constraints demonstrated in (2) above), then the degradation should follow the transition patterns listed in the right-hand

side column: an object gives its relative central position to a new object (8a), or a subject to a new subject (8b), or one of the two central arguments leaves the central zone resulting in an intransitive argument frame with the single argument necessarily bearing the Nominative in the accusative type of languages (8c–d); and finally the case illustrated in (8e) below is the "degenerate" transition variant, in which there is simply no change in argument structure (here the fact of derivation can be indicated on the basis of changes in morphological and/or semantic factors).

(8) (a) ken némi vajat a kenyérre → megkeni a kenyeret
spread some butter-acc the bread-sub → perf-spread the bread-acc
vajjal +N → +C +C → N
butter-inst
'spread some butter on the bread' → 'spread the bread with butter'

(b) Péter ás egy gödröt → Mari ásat egy gödröt
Peter dig a hole-acc → Mary dig-cause a hole-acc
Péterrel −N → −C −C → N
Peter-inst
'Peter digs a hole.' → 'Mary makes Peter dig a hole.'

(c) olvas egy regényt → olvas −C → −/+C +C → N
read a novel-acc → read
'he is reading a novel' → 'he is reading'

(d) Péter elkeni a pacát → a paca elkenődik +C → −/+C +C → N
Peter smudge the ink-blot-acc → the ink-blot smudge-middle
'Peter smudges the ink-blot' → 'the ink-blot gets smudged'

(e) Péter kavarja a levest → Péter kavargatja a levest ∅ ∅
Peter stir the soup-acc → Peter stir-dim the soup-acc
'Peter is stirring the soup' → 'Peter keeps on stirring the soup (occasionally)'

(9) (a) 1. Péter telefonál → Péter megtelefonálja a hírt +N → +C
Peter telephone → Peter perf-telephone the news-acc
'Peter is telephoning' → 'Peter phones sy about the news'

2. fakad a genny a kelésből → kifakad a kelés +N → −/+C
burst the pus the core-ela → out-burst the core
'pus bursts from the core' → 'the core bursts'

(b) 1. Péter dolgozik → Mari dolgoztatja Pétert −N → −C
Peter work → Mary work-cause Peter-acc
'Peter works' → 'Mary makes Peter work'

2. → Mari új varrónővel dolgoztat −N → −/+C
Mary new needlewoman work-cause
'Mary makes a new needlewoman work [M. has a new needlewoman]'

(c–d) 1. Péter berúg → Péter be van rúgva −/+C → −/+C
Peter into-kick → Peter into is kick-adv
'Peter gets drunk' → 'Peter is drunk'

2. érkezik három vendég → vendég érkezik −/+C → N
arrive three guest → guest arrive
'three guests arrive' → 'guests arrive [or a guest arrives]'

(e) Péter borozik → Péter borozgat ∅ ∅
Peter drink-wine → Peter drink-wine-dim
'Peter drinks wine' → 'Peter takes a glass or two of wine'

Then, in the series of examples in (9) above, the variants with an intransitive input corresponding to the five cases listed in (8) have been gone through. It can be checked that—theoretically—two "degenerate" transitions correspond to each transitive basic case; and, witnessed by the examples, what has been predicted **does** exist in Hungarian, indeed. In type (a), a positive central role is requested by an argument coming from a non-central status (8a), (9a). When the input is intransitive (9a), this request can be satisfied in two ways: the given argument either acquires a positive pole (and only this role, doing with the object position), or it deprives the single input central argument of both central roles. Variants of (9b) show the same in a symmetrical configuration: a new argument appears and acquires the subject position, while either letting the input central argument retain its object status or entirely pushing it out from the central zone. Types (8c) and (8d) coincide in the case of an intransitive input (9c–d), but the single central participant can have two fates: either a double central role can be attributed to it—resulting in an **identical** transition (which is an approach that I will argue for in section **5**), or the single argument can be deleted from the central zone, entirely emptying this zone (see also section **5**). Finally, (9e) shows that the "degenerate" variant of the identical transition can also take place when the input is "degenerate," too, that is, intransitive—which is no surprise.

In the following three sections, we will look through the Hungarian derivations qualified as productive or at least semi-productive by means of the chart in the Appendix with the purpose of demonstrating that the output case frame is determined in each case by the output **word category** and its classification according to the **five basic types of transition** discussed above; the degenerate types with an **intransitive input** can be unequivocally classified on the basis of the corresponding basic transi-

tive types, as operations of derivation (in Hungarian) will not sort out according to the cardinality of the central zone of the input

3. Kinds of derivation retaining argument structure

Let us start studying the systematizing chart in the Appendix with the first one out of the five major sections, where we are not to count with advancing or degrading arguments but only formal changes in the central zone triggered by a (potential) change in word category.

Even this latter change is minimal in the left upper square, where the output case frame practically coincides with the input case frame because of retaining the essentially verbal character. The most identical transition is characteristic of the (diminutive/frequentative) derivational suffix *-gAt*, which produces no change in category at all: it converts a verb into a verb, the input subject remains a subject, and the input object also retains its object position (see (8e) above). As indicated by a symbol "♣" in the chart, the possibility of a degenerate intransitive input is not excluded either (9e). The same could be said about the mood suffix *-hAt*; but it is so productive and its meaning contribution is so compositional that nowadays it is not considered to be a derivational suffix.

Ways of **participle** formation retaining argument structure were illustrated in (6c-e) above. The subtypes can be characterized by different slight changes in case frame as follows: the input Nominative is substituted by a possessor-like form (6e), or a form with no phonetic realization (6c); although there are special constructions (6d) where the Nominative is retained (NOM → NOM / POSS / ∅). There are three sorts of adjectival participles in Hungarian, of which the present ("continuous") participle is derived by an argument-structure retaining ASV transition; furthermore, it is often mentioned (e.g., Laczkó 2000a) that the past ("perfect") participle has a growing variant typical of the "newspaper language" (e.g., %*a tegnap játszott csapat* [the yesterday play-pastprt team] 'the team that played yesterday'), which also belongs to the type in which argument structure is retained (it expresses temporal antecedence in contrast to the present participle expressing simultaneity; cf. the normal past participle produces a passivization-like ASV transition, see section **5**).

There are three derivations forming verbs from adjectives which can be regarded as representatives of the transition retaining argument structure. At the input practically a subject should be counted with, which will undergo no change: e.g., *sötét a haja* [dark the hair-poss3sg] 'his

hair is dark' → *sötétedik a haja* [dark-become the hair-poss3sg] 'his hair is darkening', *zöld a rét* [green the field] 'the field is green' → *kizöldül a rét* [out-green-become the field] 'the field greens out', *sárga a gólyahír* [yellow the cowslip] 'cowslips are yellow'→ *sárgállik a gólyahír* [yellow-*llik* the cowslip] 'cowslips gleam yellow'. It would go beyond the scope of this paper to provide semantic changes associated with formal changes in argument structure in general but here I give an illustration of semantics: in the first two examples a static statement concerning the state of an argument x is substituted for a dynamic eventuality structure describing the change of x, whose **result state** (Alberti–Ohnmacht 2005) just coincides with (the cumulative phase of) the input eventuality structure; whereas in the last example the input static description is set in an **intensional** dimension (Alberti–Ohnmacht *ibid.*): 'it is the speaker's intensive impression that (the) cowslips are yellow'. Other sorts of semantic changes can be observed in cases of verb formation from nouns: e.g., 'x is a bashaw' → 'x behaves as if it were true that [x is a bashaw]' (*basa* → *basáskodik*), 'x is a soldier' → 'x behaves in an appropriate way in the state that [x is a soldier]' (*katona* → *katonáskodik*). What is relevant to us here: the input intransitive case frame will undergo no formal changes.

Now let us look at the argument-structure retaining subtypes of derivation with no category change: the verbal *-gAt*, which has already been mentioned, the adjectival *-(A)s*, and the family of diminutive suffixes of nouns (e.g., *-(cs)kA*). It is interesting that they share a semantic element, some diminutive character, pervading the borderline of word categories: 'he *does* that occasionally/not so seriously'/'it is essentially *such* but the given property does not manifest itself in its entirety'/'it is *that* but smaller/less developed than the prototypical version'.

The adjectival privative suffix semantically produces an opposite property ('x is brave' → 'x is not brave'), but the case frame will not change (x_{NOM} *bátor/bátortalan*).

As for the suffix *-An*, I follow Kiefer and Ladányi (2000b, 4.4.1.) in considering it a derivational morpheme forming adverbs; in this approach it is obviously a transition retaining argument structure: e.g., *részeg* → *részegen* 'x is drunk' → 'x does something whilst [x is drunk]'.

The derivations mentioned in the chart turning nouns into adjectives (e.g., *király* 'king' → 'super' (in slang), *gáz* 'gas'→ 'unpleasant' (in slang)) and adjectives into nouns (e.g., *angol* 'English' → 'Englishman') also retain argument structure. It should be noted in connection with the type of nouns like *autószerelő* [car-repair-presprt] 'car-mechanic' (noun) that

the immediate source of this noun is an adjectival version of *autószerelő* (*autószerelő brigád* 'team repairing cars'), to be formed by conversion, and this adjective is formed from the present participial construction *autót szerelő* [car-acc repair-presprt] 'one repairing cars' through a kind of derivation which will be mentioned as **patient incorporation** in section **5**.

I follow Laczkó (2000b) in classifying *-ÓjA* [presprt+poss] as a non-composite suffix forming nouns from verbs; and in the extended version of Model Tau we can claim that it is just this derivational morpheme that realizes noun formation from verbs in an argument-structure retaining way. Remember the last section: what corresponds to the verbal central case frame ⟨NOM, ACC⟩ in the category of nouns is a central frame ⟨NOM, POSS⟩. The examples below, thus, show instances of identical, but intercategorial, transition:

(10) (a) Mari szereti / elcsábítja Pétert. ⟨NOM, ACC⟩ = ⟨−, +⟩
Mary love-3sg / seduce-3sg Peter-acc
'Mary loves/seduces Peter.'

(8e)

(b) Mari a szeretője / elcsábítója Péternek. ⟨NOM, POSS⟩ = ⟨−, +⟩
Mary the love-*Ója* / seduce-*Ója* Peter-dat
'Mary is Peter's sweetheart/seducer.'

4. Kinds of derivation expanding argument structure

As was elucidated in section **1**, Model Tau relies on the hypothesis (also of distinguished importance in Zsilka's (1966, 1982) philosophy) that languages can get accustomed to the changing world via the development of **polarized chains of influence**; and the creative element of this process is nothing else but the two sorts of ASV transition **expanding** argument structure.

First of all, let us consider the derivational morphemes mentioned in the chart which enrich the chain of influence on the side of "causees". It has been discussed in connection with (8a) above that what practically happens in the basic case of this type of transition (where the input is transitive) is that the input object is substituted by a new argument in the object function. Verbs *ken* 'spread' and *fakaszt* 'cause to burst' are representatives of two distinct basic types. In the former case the content of the transition can be characterized as follows: in the input situation an Agent moves a Patient to a Goal (e.g., 'somebody spreads some butter on a bread') while in the output situation our attention is concentrated

on the fact that the Goal has been affected (totally) as a result of this movement ('he spreads the bread with butter'). In the latter case the input situation describes a movement in the opposite direction, that is, the Patient moves away from a Source (e.g., 'he has some pus burst from the core'), while in the output situation we concentrate our attention on the change of state pertaining to the Source, i.e., its becoming empty ('he has the core burst').

The two sorts of transition discussed in the last paragraph do not require an Agent to move the Patient: we can speak about the total affectedness of a Goal by the saturation characteristic of it (*zsonganak a méhek a kertben* [swarm-3pl the bee-pl the garden-ine] 'bees are swarming in the garden' → *zsong a kert a méhektől* [swarm-3sg the garden the bee-pl-abl] 'the garden is swarming with bees'; *sárgállanak a gólyahírek a réten* [yellow-*llik*-3pl the cowslip-pl the field-sup] 'cowslips gleam yellow in the field' → *sárgállik a rét a rengeteg gólyahírtől* [yellow-*llik*-3sg the field the many cowslip-abl] 'the field is yellow with the many cowslips), and the total affectedness of a Source by the emptiness characteristic of it (*fakad a genny a kelésből* [burst the pus the core-ela] 'pus bursts from the core' → *kifakad a kelés* [out-burst-3sg the core] 'the core bursts'). The degenerate intransitive input follows the transition pattern demonstrated in (9a.2): the argument playing the positive central role in the input ASV leaves the central zone, so the new argument entering the central zone promptly obtains a double polarity, and the resulting output is an intransitive ASV again.

In the case of an intransitive input, the other transition pattern predicted in (9a.1) is also quite frequent: in this pattern the single central argument of the input ASV gives the new argument of the central zone only its positive central feature, while retaining the negative feature. What is produced in this way is a transitive output version: e.g., *telefonál* 'telephone' → *megtelefonálja a hírt* [perf-telephone-3sg the news-acc] 'phone sy about the news'. The existence of the new argument is due to the fact that the sentences "emitted" in the course of telephoning have constituted a coherent unit that can be characterized as a piece of news; thus the Agent's permanent activity which does not necessarily aim at a purpose in the input situation is represented as a telic event in the output version: the purpose is producing a Patient like in the transition type of *fon* 'weave' (Hungarian shows the same transition: 'weave the silk thread into a shirt' → 'weave a shirt').

Transition pattern (9a.1) is a very productive source of ASVs often described as constructions with "pseudo-objects", in which the object does not belong to the verb as its argument ("szótárilag nem vonzata az igének"; e.g., Bene 2005). The activity of swimming, for instance, is basically expressed by means of an intransitive ASV (see (11a) below), but numerous transitive versions can be formed (11b–f):

(11) (a) Péter úszik.
Peter swim-3sg
'Peter is swimming.'

(b) Péter (le)úszott öt kilométert.
Peter (down-)swim-past-3sg five km-pl-acc
'Peter has swum five kilometers.'

(c) Péter átúszta a Csatornát.
Peter across-swim-past-3sg the Channel-acc
'Peter has swum the Channel.'

(d) Péter végigúszta a délutánt.
Peter throughout-swim-past-3sg the afternoon-acc
'Peter swam all afternoon.'

(e) Péter világcsúcsot úszott.
Peter world-record-acc swim-past-3sg
'Peter swam a world record.'

(f) Péter leúszta Pált.
Peter down-swim-past-3sg Paul-acc
'Peter and Paul swam a race and Peter won.'

(g) Péter összeúszott magának egy nyaralót.
Peter together-swim-past-3sg self-dat a cottage-acc
'Peter has earned so much money by swimming professionally that he could buy a cottage.'

(h) Péter halálra úszta magát.
Peter death-sub swim-past-3sg self-acc
'Peter had swum himself to death.'

The following question arises in connection with Bene's (2005) point of view: how can we know what belongs to a verb "lexically" as its argument? The approach based on pseudo-objects is problematic because the completion of an event can be measured out by considering just the participant in question, which is characteristic of arguments, moreover, of distinguished arguments (Tenny 1994). If somebody intends to swim

five kilometers, for instance, and he has begun to do that, then at a given moment it can be measured out according to the distance absolved up to that particular moment how matters stand, and the entire event can be considered to have been accomplished successfully (i.e., it can be said that he has reached the cumulative point (Alberti–Ohnmacht 2005) when he has absolved the fifth kilometer.

It is quite clear why "pseudo-objects" cause problems to traditional theories of thematic roles: a threatening number of ASVs can be produced with them, different sorts, as could be seen above, but fairly independently of the idiosyncratic properties of verb stems, and in all these ASVs there seems to be no better approach then labelling them as Patients, which is no favorable result for an ambitious thematic theory. In Model Tau, however, in which thematic role labels have been got rid of and more abstract argument hierarchies are relied on (see section 1), nothing prevents us from proposing what is suggested by the appearance of the Accusative case: they are ordinary arguments occupying the object position. We hypothesize that **alternative** chains of influence develop with arguments **incommensurate** according to the partial ordering that the sum of the family of these chains of influence constitutes, as in the case of *ás* 'dig' (see also (1)):

(12)	a.	Péter	→	föld	'ground'	→	kert	'garden'	(*felás* [up-dig])
	b.	Péter	→	föld	'ground'	→	gödör	'hole'	
	c.	Péter	→	föld	'ground'	→	gödör	'hole'	→ sír 'grave'
	d.	Péter	→	föld	'ground'	→	gödör	'hole'	→ kincs 'treasure' (*elás* [away-dig])
	e.	Péter	→	föld	'ground'	→	gödör	'hole'	$\rightarrow_{\text{out}}$ kincs 'treasure' (*kiás* [out-dig])

As can be seen above, it can be expressed in Hungarian by using an appropriate object (and an appropriate verbal prefix in certain cases) that a whole garden is affected by digging its soil (12a), whilst no real holes are created; in other ASVs (12b–e) a real hole is created, which can serve as a grave (12c) or a place to hide a treasure in it (12d–e); and it depends on the morpheme appearing as a prefix of the verb stem *ás* 'dig' whether the treasure gets in this place (12d) or out of this place (12e).

Thus the intricate system of relations between the numerous arguments "swarming" around a given verb stem can be accounted for by means of the partially ordered structure of families of branching chains of influence. As for the calculation of the meaning of a particular ASV,

what is to be allowed for, in addition to the verbal prefix if any, is the semantic type of the object (e.g., time, distance, place, person). The interpreter's task is practically to figure out the likeliest potentially **relevant** relation between a person swimming and a time mentioned, or a place, or a person, or the Agent himself/herself (11b–h) (in a sufficiently richly structured lexical network, Alberti 2000). It is a secondary question whether a given meaning should be calculated from its pieces again and again on the basis of the lexical network, or registered as a new item.

Now let us turn to the mysterious suffix *-((j)A)*, which appears in the word *kalap-ja-i-m* [hat-poss-pl-1sg] 'my hats' in its full-fledged form but seems to be present in a null phonetic form in *hajó-Ø-i-m* [ship-poss-pl-1sg] 'my ships'. We follow Szabolcsi (1992) and her predecessors in considering this morpheme to indicate the state of being possessed. I would like to add, however, the uncustomary hypothesis that this suffix should be regarded as belonging to the derivational elements considering its property of increasing the number of arguments: e.g., *x kalap* '*x* is claimed to be a hat' → *x Péter kalapja* '*x* is claimed to be Peter's hat', i.e., a hat and Peter are claimed to stand in a certain relation (cf. **x Péter kalap*). As a derivational morpheme, it should be placed in the cell of nominal derivational elements coming with no category change which realize the transition pattern demonstrated in (9a.1) with an "intransitive" input (where 'intransitive' in the nominal category is to mean that a noun's ASV contains no possessor):[4]

(13) (a) Ez itt (egy) kalap. ⟨NOM⟩ = ⟨−/+⟩
this here (a) hat
'This here is a hat.'

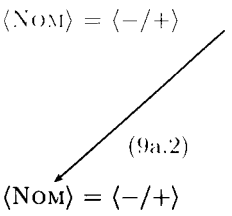
(9a.1)

(b) Ez itt a kalapja az én feledékeny ⟨NOM, POSS⟩ = ⟨−/+⟩
this here the hat-poss3sg the I absent-minded
Péter barátomnak.
Peter friend-poss1sg-dat
'This is the hat of my absent-minded friend Peter.'

[4] The other nominal derivative suffix *-ÓjA* (Laczkó 2000b), discussed in section **4**, obviously contains *-((j)A)* (accompanied by the suffix of present participles), which we claim to be another argument in favor of classifying this (latter) suffix as a derivational element.

(14) (a) Gyakran darálnak a munkások ebben
often grind-3pl the worker-pl this-ine
a műhelyben.
the workshop-ine
'The workers often grind in this workshop.'

(b) Ez itt a daráló.
this here the grind-presprt
'This is the grinding workshop.'



In (14) above I have illustrated the derivational suffix *-Ó* proposed by Laczkó (2000b) as a morpheme forming names of places following transition pattern (9a.2) (which is a transition with an intransitive input). An analogous example in the area of V → V derivations is *épülnek a házak a hegyoldalon* [build-3pl the house-pl the hill-side-sup] 'houses are built on the hill-side' → *beépül a hegyoldal* [into-build-past-3sg the hill-side] cca. 'the hill-side has been covered by new buildings'—which is an analogy that we judge to corroborate the hypothesis on the suffix *-Ó* because of the similarity in respect of meaning.

Let us turn to the other subtype of argument-structure expanding derivations, in the course of which the chain of influence enriches at its negative edge. Let **causative/factitive derivation** be the first ASV transition to be commented on. As was shown by (8b) in section **2**, when the input is transitive, there is a replacement in the position of subject with object position remaining unaffected. In the (degenerate) case of an intransitive input, it is (9b.1) that can be called the productive transition pattern, in which the new subject "pushes" the input subject into the object position. As for the other intransitive transition pattern, shown in (9b.2), a few examples can be mentioned, which we judge to be relics in the synchronic state of language. In this latter subtype the output ASV is also intransitive (e.g., *Hol/Kivel varratsz/lektoráltatsz mostanában?* [where/who-inst sew-cause-2sg/revise-cause-2sg nowadays?] 'who is your new tailor/literary adviser?').

In connection with causative derivation it is worth discussing the distinction of the "phonetic reality" of particular derivational morphemes from the abstract level of ASV transition belonging to them. As for the former aspect, only *-(t)At* is claimed to be productive in Hungarian (Komlósy 2000) (e.g., *dolgoztat* [work-cause] 'make sy work', *ásat* [dig-cause] 'make sy dig'; *-Aszt*, for instance, is not productive (but see *fagyaszt* [freeze-*Aszt*] 'make sth frozen'). Word form **fagy(t)at*, how-

ever, does not exist in Hungarian. Its derivation is blocked, which can be attributed to the following mechanism: there is a transition *fagy* → *fagyaszt* in Hungarian which can be regarded as a relic that has survived in the synchronic state of the language, and this relic makes it redundant, and hence forbidden, to apply the productive procedure of derivation. A uniform picture of abstract transition patterns and phonetic variants can be worked out by understanding their relation as follows: the primary factor is the patterns of ASV transitions belonging to certain cells of the chart in the Appendix and associated with a predictable change of meaning, and it should be studied then what kind of phonetic form **realizes** certain types of transition depending on particular types of inputs. The picture may be fairly intricate in certain areas: in the area of Hungarian causative/factitive derivation, for instance, one productive suffixation works (*-(t)At*), which "keeps away", on the one hand, from transitions which are surviving relics (e.g., *fő(l)* → *főz* 'be on the boil' → 'cook', *süllyed* → *süllyeszt* 'sink$_{intr}$' → 'cause to sink'), and, on the other hand, from input–output pairs produced by other derivations (e.g., *sárgul* → *sárgít*, but **sárgultat* 'become/make yellow', *sötétedik* → *sötétít*/**sötéted(t)et* 'become/make dark'). The possibility of distinguishing these two levels is an advantageous feature of Model Tau. Transition patterns can be defined without referring to morphemes of derivation: they are elements of UG, which are expressed in a given language by means of diverse morphological tools.

Now let us scrutinize the derivational suffix *-ít*, mentioned above. Its input case frame is necessarily intransitive because of the adjectival category, but the verbal category in the output already makes a transitive frame possible, which is due to pattern (9b.1). It is worth making a comparison between the productive *-ít* and the improductive *-Vll*, which follows also transition pattern (9b.1). The essence of their difference can be elucidated by referring to thematic roles (but should be captured and calculated in Model Tau in another way): *-ít* enriches the chain of influence with an Agent while *-Vll* with an Experiencer (*drágít* [expensive-*ít*] 'make sth (more) expensive', versus *drágáll* [expensive-*áll*] 'consider sth to be expensive').

Examples (15)–(17) below serve as an illustration of relevant cases of noun formation. The corresponding transition patterns have also been given below. In (15) and (17) the argument obtaining the subject position—independently of the verbal, adjectival or nominal category of the input word—will be the **Davidsonian** or **eventual** argument of the input

situation (e.g., Parsons 1995): something (the dangerous or surprising situation) is predicated of that it is nothing else but the **activity** of a lion's stroking or the **fact** that Mary is beautiful or a teacher.[5] Type (15) has a transitive input, and in the subject position the Agent of the input will be replaced in the output ASV with what serves as the eventual argument in the input. Transition type (17) is degenerate as its input is intransitive; the transition can be described by (9b.1): the input subject will accept a possessor's function in order to be able to give the subject status to the eventual argument. The type demonstrated in (16) below (e.g., *a kutya harapása* [the dog bite-nominal-poss3sg] 'the bite of the dog', which is a **result**, and not an **activity/state** (see above)) is characterized by strictly restricted productivity, so I mention it only because of the famous ambiguity shown in (15)/(16). The transition in (16) can be captured essentially in the same way as those demonstrated in (17), with intransitive inputs (*a kutya harap* 'the dog bites').

(15) (a) Mari simogatja az oroszlánt. ⟨NOM, ACC⟩ = ⟨−, +⟩
Mary stroke-3sg the lion-acc
'Mary is stroking the lion.'

(8b)

(b) Ami igazán veszélyes, az az oroszlánnak$_{PAT}$ ⟨NOM, POSS⟩ = ⟨−, +⟩
what really dangerous that the lion-dat
a simogatása.
the stroke-nominal-poss3sg
'What is really dangerous is the stroking of the lion.'

(16) (a) ?Az oroszlán$_{AG}$ ritkán simogat (inkább üt ⟨NOM⟩ = ⟨−/+⟩
the lion rarely stroke (rather beat
és karmol).
and scratch)
'A lion rarely strokes, it rather beats and scratches.'

(9b.1)

(b) Amit itt láthatsz a karomon ⟨NOM, POSS⟩ = ⟨−, +⟩
what-acc here see-mod-2sg the arm-poss-1sg-sup
az az oroszlánnak$_{AG}$ a simogatása.
that the lion-dat the stroke-nominal-poss3sg
'What you can see here on my arm is the lion's stroking.'

[5] Sentence (15a), for instance, is about three participants: the two "normal" arguments of the verbal predicator, Mary and a lion, and the fact that Mary is stroking the lion, which can be regarded as an additional, or rather, "the 0-th," argument. An independent piece of evidence in defence of this approach is the

(17) (a) Mari szép / tanár. ⟨NOM⟩ = ⟨−/+⟩
Mary beautiful / teacher
'Mary is beatiful / a teacher.'

(9b.1)

(b) Ami igazán meglepett, az Mari ⟨NOM, POSS⟩ = ⟨−, +⟩
what really surprise-past-3sg that Mary
szépsége / tanársága.
beautiful-nominal-poss3sg / teacher-nominal-poss3sg
'What has really surprised me is the fact that Mary is beautiful/a teacher.'

(18) (a) Ez itt Pécs. ⟨NOM⟩ = ⟨−/+⟩
this here Pécs
'This here is Pécs.'

(9b.2)

(b) Mari pécsi. ⟨NOM⟩ = ⟨−/+⟩
Mary Pécs-adj
'Mary is from Pécs.'

Finally a derivational suffix productively turning a certain type of nouns into adjectives should be mentioned, whose phonetic form is *-i*. As is illustrated by (18) above, now we should have recourse to the pattern of transition (9b.2), pertaining to intransitive inputs, which replaces a single central subject with another one.

5. Kinds of derivation reducing argument structure

This section is devoted to commenting on the last two parts of the systematizing chart in the Appendix, which contain ASV transitions essentially deleting one of the central arguments, at least from the central zone.

A transition like this can also be understood as the advancement of a central argument by assigning both central poles to it. As was pointed out in connection with (8) in section **2**, advancement and degradation mutually trigger each other; a derivational transition, hence, can be attributed to one of these two operations, which can be called the **decisive factor** of the given transition, whilst qualifying the other operation as an additional element. It will be pointed out that the decisive factor can be

observation that these three participants can be referred to by (different sorts of) pronouns: e.g., *Mary*$_1$ *is stroking the lion*$_2$ *though she*$_1$ *can see that it*$_2$ *does not like that*$_0$. *Mary*$_1$, *who*$_1$ *is my brother's daughter, is stroking a lion*$_2$, *which*$_2$ *arrived at the zoo yesterday, which*$_0$ *requires serious braveness.*

decided on unambiguously if a derivation operates on intransitive inputs as well. In the chart the cell of each transition type is divided into two parts horizontally, and I show my hypothesis on the decisive factor of the given transition (advancement/degradation). I will note where the decision is obvious.

Let us start the detailed discussion with derivations deleting a **positive** central argument. In the simplest case an object is deleted, which can be called **intransitivization** and is realized in the form of conversion (in Hungarian), with limited productivity (*Péter eszik/*helyez* 'Peter eats/ *places'). The change in meaning can be approximated by the following formula: $\text{ESZIK}_{\text{intr}}(x) = \exists y.\text{ESZIK}_{\text{tr}}(x, y)$. In this logical formula, argument y is existentially bound, so syntax can provide for it no argument slot that could be freely occupied any more. The reason of this kind of semantic modification may lie in various factors: we cannot, or do not intend to, specify the kind of food (in the given case), or we intend to raise the hearer's attention to the Agent's impact upon himself/herself, i.e., (s)he is busy with eating or (s)he makes himself/herself strong or fat.

A syntactic argument—as a slot that can be filled in freely—can be expired also by identifying two arguments: for example, $\text{FÉSÜL}_{\text{refl}}(x) = \text{FÉSÜL}_{\text{tr}}(x, x)$ ('comb'/'comb oneself'). This kind of transition can be called **reflexivization** and is realized in Hungarian by a family of similar suffices (e.g., *fésülködik, mosakodik, borotválkozik* 'comb/wash/shave oneself'). Reflexivization makes it explicit that the Agent exerts some kind of influence upon himself/herself.

The object argument slot that can be filled in freely can also be ceased by **incorporation**. This transition via conversion can be characterized as follows: the object which typically appears in a non-determined form moves to the place immediately preceding the verb stem (while the sentence remains neutral) and forms a word-size intonational unit with it (with a single stress on the first syllable—in accordance with Hungarian phonology) (Komlósy 1992). The following formula is an illustration of the change in meaning: $\text{SZEREL}_{\text{inc:AUTÓ}}(x) = y.[\text{SZEREL}_{\text{tr}}(x, y)\ \&\ \text{AUTÓ}(y)]$ ('repair' → 'car-acc+repair', see (19a–b) below). As is shown in (19b) below, what can be incorporated can be characterized as an object which is **typical** relative to the event, i.e., with which the given event can be regarded as "institutionalized" in a generalized sense (Komlósy *ibid.*). In (19a–b) the **identity** of the three types of transitions can be observed; what is different is the change in meaning, see the formulas in (19c.1–3) below:

(19) (a) Péter eszik egy zsemlét / megfésüli magát / ⟨Nom, Acc⟩ = ⟨−, +⟩
Peter eat-3sg a roll-acc / perf-comb-3sg self-acc /
szereli az autót.
repair-3sg the car-acc
'Peter eats a roll/combs himself/repairs the car.' (8c)

(b) Péter eszik / fésülködik / autót / ⟨Nom⟩ = ⟨−/+⟩
Peter eat-3sg / perf-comb-refl-3sg / car-acc /
???Fordot szerel.
Ford-acc repair-3sg
'Peter eats/combs himself/repairs cars/Fords.'

(c) 1. $V_{intr}(x) = \exists y.V_{tr}(x, y)$
2. $V_{refl}(x) = V_{tr}(x, x)$
3. $V_{inc:N}(x) = \exists y.[V_{tr}(x, y)\ \&\ N(y)]$
4. $V_{inc:N}(\ ) = \exists y.[V_{intr}(y)\ \&\ N(y)]$

Is it possible in the case of the three sorts of derivations discussed above that the input is degenerate, i.e., intransitive? The answer is positive in the case of incorporation (and only in this case), as is witnessed by the example in (20) below: it has been observed that a Patient can be incorporated without the presence of an Agent in the argument structure whereas a lonely Agent cannot be incorporated. See the formula in (19c.4) above: the central zone of the output ASV in this degenerate case is empty. This circumstance also reveals that the decisive factor of incorporation is the degradation of the argument bearing the positive central role (and not the Agent's advancement). Thus **positive** central arguments will undergo this operation; and what (20b) shows is that the same can also be claimed by referring to the **absolute** character: independently of the input number of arguments, essentially Patients are pertained to (providing an example of Perlmutter's (1978) **Unaccusative Hyposthesis**).

This observation can be captured in Model Tau as follows: if the **decisive factor** of a transition in the case of which also an intransitive input is permitted is the operation (argument degradation/advancement) pertaining to the **positive** central argument, then, out of the arguments ordered in the chain of influence, there is an argument in the primitive core with the following property: the given derivation can be applied to the arguments in the chain in the positive direction, but not in the negative direction, relative to this distinguished argument. A strong agent, thus, cannot be incorporated (**régész ásat itt* 'archeologist dig-cause here'), whereas in the case of a strong patient incorporation is

theoretically permitted and its realization will depend on the pragmatic criterion of institutionalization, mentioned above (*gödröt/sírt/aranyat ás* 'hole-acc/grave-acc/gold-acc dig').

(20) (a) Érkezett / Telefonált három finn vendég. ⟨Nom⟩ = ⟨−/+⟩
arrive-past-3sg / phone-past-3sg three Finnish guest
'Three Finnish guests arrived/telephoned.'
(9c.2)
(b) Vendég érkezett / *telefonált. ⟨ ⟩ = ⟨∅⟩
guest arrive-past-3sg / phone-past-3sg
intended meaning: 'One or more guests arrived/telephoned.'

Let us return to cases of object degradation with a transitive input in favor of the versions of suffix *-Ó* forming adjectives (and not present participles), carefully classified by Laczkó (2000a). Combining incorporation with adjective formation can result in the type of the example *autószerelő (brigád)* ['car-repair-*Ó* (team)'] 'team repairing cars' (which can also be regarded as the first step towards the noun *autószerelő* ['car-repair-*Ó*'] 'car-mechanic', whose formation requires a second step discussed earlier, viz. a conversion retaining argument structure). In the type of adjective formation exemplified by *a szerető*$_{\mathrm{A}}$ *fiad* ['the love-*Ó* son-poss2sg'] 'your son who loves you', the object cannot be chosen freely (cf. *a Marit szerető*$_{\mathrm{V}\rightarrow\mathrm{A}}$ *fiad* ['the Mary-acc love-*Ó* son-poss2sg'] 'your son who loves Mary'), so it has been deleted as a syntactic argument.

In (21)–(22) below analyses concerning noun formation are demonstrated. (21a) is intended to show (by means of a fictive verb stem) that *-Ó* as a suffix forming nouns is productive in two versions, with output words referring to actors/instruments (Laczkó 2000b). Transition pattern (8c) accounts for both versions, which operate on different inputs, obviously.

(21) (a) Az a munkás / gép ott álló nap
that the worker / machine there whole day
strimpfeli a *biszcájgokat.*
strimpf-3sg the *bißzeug*-pl-acc
'That worker/machine there *strimpfs* the *bißzeugs* all the day.'

(b) Ő az új strimpfelő. / Az a gép egy
he the new *strimpf*-Ó / that the machine a
új német strimpfelő.
new German *strimpf-Ó*
'He is the new *strimpfer.* / That machine is a new German *strimpfer.*'

⟨NOM, ACC⟩ = ⟨−, +⟩
(8c)
⟨NOM⟩ = ⟨−/+⟩

(22) (a) Ady új verse /
Ady new poem-poss3sg /
az autó örökös szerelése
the car unending repairing-poss3sg
'Ady's new poem' / 'the unending repairing of the car'

(b) Ady-vers / autószerelés
Ady-poem / car-repairing
'poem of Ady's' / 'repairing of cars'

⟨NOM, POSS⟩ = ⟨−, +⟩
(8c)
⟨NOM⟩ = ⟨−/+⟩

The type of noun formation retaining category, demonstrated in (22) above, can also be traced back to transition pattern (8c), as the input word category defines the ⟨NOM, POSS⟩ input case frame; what is yielded, then, as an output is a complex word including the original possessor incorporated. Note that the central frame, which is already rid of possessors, can be enriched with a possessor again in the way discussed in the previous section: *Péter(-nek a) kedvenc Ady-verse* 'Peter(-dat the) favorite Ady-poem-poss3sg' *Péter(-nek az) örökös autószerelése* 'Peter(-dat the) unending car-repairing-poss3sg'; while two possessors cannot remain in the central frame (despite the semantic possibility mentioned above), which is a correct prediction.

The last major section of the systematizing chart demonstrates transitions in the course of which the input **subject** is deleted (from the central zone), resulting in the input object advancing into a subject bearing a double central role (8d). In English, **passivization** can be characterized in this way (Alberti 1996, 1998a), whose crucial function in the system of this language is associating the Patient with a **topic** pragmatic role (closely related to the subject position in English); the cost, as a result

of constraint (2a), is that the Agent should be deleted from the central frame as it is not permitted to appear as an object in the same ASV.

In Hungarian, passivization as a sort of verb formation is archaic, and, hence, cannot be characterized as productive (e.g., *riporter kerestetik* ['reporter seek-pass'] 'reporters are looked for', *ilyen lehetőség ritkán adatik az embernek* ['shuch chance rarely give-pass the man-dat'] 'a chance like this is rarely given to you'), and the formation of *mediális* (cca. 'middle') forms comes with a special additional meaning (e.g., *magától megoldódott* ['self-abl solve-mid-past'] 'solved of its own accord'); but participial and nominal outputs can be formed in many ways according to transition pattern (8d), as is shown by the densely covered lower part of the chart. (23) shows the straightforward cases:

(23) (a) Péter megoldotta a problémát. ⟨NOM, ACC⟩ = ⟨−, +⟩
Peter perf-solve-past-3sg the problem-acc
'Peter solved the problem.'

(8d)

(b) A probléma megoldódott / ⟨NOM⟩ = ⟨−/+⟩
the problem perf-solve-mid-past-3sg /
†megoldatott / meg van oldva. /
perf-solve-mid-past-3sg / perf is solve-adv /
megoldott / megoldandó / megoldható /
perf-solve-pastprt / perf-solve-futprt / perf-solve-*hAt* /
megoldhatatlan / megoldatlan probléma
perf-solve-*hAtAtlAn* / perf-solve-*AtlAn* problem
'The problem was solved / was solved by sy / is solved.' / 'a problem which [has been solved] / [is to be solved] / [can be solved] / [cannot be solved] / [is not solved]'

Further comments are due on the sorts of derivation accepting even intransitive inputs: e.g., those forming adverbial participles (*ki van apadva* ['out is dry-adv'] 'is exhausted'), and past participles (*kiapadt* ['out-dry-pastprt'] 'exhausted'); *-hAtAtlAn* forms appear in numerous relics (*kiapadhatatlan* ['out-dry-*hAtAtlAn*'] 'inexhaustible') but this sort of derivation is not productive (cf. **kifáradhatatlan* ['out-tire-*hAtAtlAn*'] 'untiring'). The **decisive factor** of the sorts of derivation, thus, is obviously the advancement of Patient (the argument with the positive central role), and not the degradation of Agent (Alberti 1996, 1998a), as the single argument will not undergo deletion; further, just intransitive ASVs consisting of a Patient will undergo these derivations, serving as an argument in

favor of the law concerning transitions with a patient-like decisive factor, discussed above example (20).

(24) (a) Péter berúgott / odafutott Marihoz. ⟨NOM⟩ = ⟨−/+⟩
Peter into-kick-past-3sg / towards-run-past-3sg Mary-all
/A rúd elgörbült.
/the bar away-curve-past-3sg
'Peter got drunk / ran to Mary.' / 'The bar curved.'
(9d.1)

(b) Péter be van rúgva / *oda van futva Marihoz. / ⟨NOM⟩ = ⟨−/+⟩
Peter into is kick-adv / towards is run-adv Mary-all /
elgörbült rúd
away-curve-pastprt bar
'Peter is drunk / has run to Mary.' / 'bar that has curved'

One might think that the Patient in an intransitive case frame **cannot** be advanced, as it already bears a double central role in the input ASV. This is true, but the functioning of advancing/degrading operations will be disturbed by no formal factors; in the degenerate situation they function as an identical mapping, with a benefit typical of the given category change (see Alberti–Ohnmacht 2005, and also Alberti 1996, 1998a).

There is a theoretical alternative: intransitive passivization might follow transition pattern (9d.2). The German *Zustandpassiv*, for instance, works in this way: e.g., *Hier ist bis 23 Uhr getanzt* ['hier is until 23 o'clock perf-dance-pastprt'] 'one is allowed to dance here until 23.00'. As is pointed out by Tóth (2000), there is a narrow but clear-cut area (**household** verbs) within which adverbial participles can be formed also in Hungarian in this way (by deleting the single argument as a decisive factor of transition where this single argument should be agent-like):

(25) (a) Mari kitakarított / beágyazott a hatosban. ⟨NOM⟩ = ⟨−/+⟩
Mary out-clean-past-3sg / into-bed-past-3sg in six-ine
'Mary did room 6/turned down beds in room 6.'
(9d.2)

(b) A hatosban ki van takarítva / be van ágyazva. ⟨ ⟩ = ⟨−/+⟩
the six-ine out is clean-adv / into is bed-adv
'Someone did room 6 / turned down beds in room 6.'

Let us return to the regular transition pattern provided in (8d). We can use this pattern to capture three types of derivation relying on properties of the central case frame attributed to nouns; see (26)–(28) below. The analysis shown in (26) provides a formal rule describing derivations with

a nominal central frame as the starting point, i.e., essentially a possessive construction. Who is *gondos* ['care-*(V)s*'] 'careful', for instance, takes *care* of the things (s)he is responsible for: the analysis in (26) below relies on this connection. An interesting question arises here: the ambiguity between *gondtalan/gondatlan* 'carefree'/'careless' (in which the same stem is furnished with two variants of the privative suffix *-tAlAn/(A)tlAn*), which can be explained as follows: *gond* 'care' can be understood as both a positive concept ('attention') and a negative one ('problem'), and this alternative in addition to the alternative forms of the privative suffix has made it possible to develop two different adjectival meanings in the course of the evolution of Hungarian. The transition pattern itself, after that, can be assumed to be the same. The derivational suffix *-(j)Ú* is special: while deleting the input subject (or rather, embedding it in the phonetic form of the predicator itself), it also has an additional effect outside the central zone, viz., it produces a predicative argument of category adjective (which corresponds to a qualifier in the input, which is a free adjunct) (Laczkó 2000c). This specialty, however, does not disturb our claims concerning changes in the central case frame.

(26) (a) Ez Marinak a gondja / (hosszú) haja. ⟨NOM, POSS⟩ = ⟨−, +⟩
this Mary-dat the care-poss3sg / (long) hair-poss3sg
'This is Mary's problem/(long) hair.'
(8d)

(b) Mari gondos / gondtalan / gondatlan / ⟨NOM⟩ = ⟨−/+⟩
Mary care-*(V)s* / care-*tAlAn* / care-*(A)tlAn* /
hosszú hajú.
long hair-*(j)Ú*
'Mary is careful / careless / care-free / long-haired.' (8d)

(27) (a) Péter felrakta az árut. ⟨NOM, ACC⟩ = ⟨−, +⟩
Peter up-load-past-3sg the commodity-acc
'Peter loaded the commodity (onto sth).'
(8d)

(b) Ez az áru viszonylag kellemes rakomány. ⟨NOM⟩ = ⟨−/+⟩
this the commodity fairly pleasant load
'This commodity is a fairly pleasant load.'

The analysis above in (27) provides the transition pattern belonging to suffix *-(V)mÁny*, which is very frequent but does not qualify as a morpheme of a productive derivation forming nouns (Kiefer–Ladányi 2000b). This is the transition pattern shown in (8d); and we should have recourse to its degenerate variant (9d.1) in favor of such examples with an intran-

sitive input (consisting of a Patient!) as *fejlemény* ('develop/-ments'), *eredmény* ('grow out' → 'outgrowth'), *esemény* ('occur/-rence').

Finally let us consider the analysis requiring the most imagination. The derivational suffix *-l/-z* forms verbs from nouns, so our starting point should be a possessive construction:

(28) (a) Ez Péternek a kapája / gereblyéje / ⟨NOM, ACC⟩ = ⟨−, +⟩
this Peter-dat the hoe-poss3sg / rake-poss3sg /
gitárja / csaja.
guitar-poss3sg / girl-poss3sg
'This is Peter's hoe/rake/guitar/girl-friend.'

(8d)

(b) Péter kapál / gereblyéz / ⟨NOM⟩ = ⟨−/+⟩
Peter hoe-3sg / rake-3sg /
gitározik / csajozik.
play-the-guitar-3sg / go-out-with-a-girl-3sg
'Peter is hoeing/raking/playing the guitar/going out with a girl.'

Is it possible to base this formal analysis upon any kind of connection between an activity and possession (interpreted appropriately)? I think so; and the key to a solution like this lies in finding the proper interpretation of the construction *X's Y* from the numerous interpretations. This input interpretation is as follows: *X-nél van Y* ('X-ade is Y') 'there is a Y with X', and then the output meaning can be formulated in this way: *X is doing something with Y that is with him/her, which is an activity straightforwardly resulting from the inherent nature of Y.* What one can do with a hoe or a guitar if properly used, for instance, is hoeing or playing the guitar.

6. Summary

Having reviewed all sorts of derivation characterized as productive in the volume on morphology of *Strukturális magyar nyelvtan* (Kiefer 2000), I claim that the *a priori* hypothesis has been corroborated: the original version of Model Tau (Alberti 1997), dealing only with verbal derivation coming with no category change, can be extended to the entire spectrum of derivations (typically coming with category change); moreover, it can be extended in the most straightforward way possible, according to which the single novel factor to be allowed for is the **central case frame** peculiar to particular word categories.

I claim that if the predicator is a noun, what corresponds to the case frame ⟨NOM, ACC⟩ in the sphere of verbs, is the case frame ⟨NOM, POSS⟩. This mapping is immediately observable in the case of *-ÓjA*, a suffix forming nouns in an argument-structure retaining way (Laczkó 2000b); see (10) in section **3**. As for the totally productive suffix *-Ás*, forming nouns with an output meaning referring to facts/activities, it is less easy to capture its impact concerning the modification of case frame (15), but the pattern of transition corresponds to nothing else but the causative derivation (8b) in the sphere of V→V formation, which is also prominently productive. Now, instead of entering into further details, I would like to highlight the essence of Model Tau's theory on derivation: (8) shows the five patterns of case-frame transitions which can be calculated as theoretical possibilities in the case of a transitive input (independently of the pair of word categories concerned), and then (9) is intended to review the degenerate cases, i.e., transitions with an intransitive input ASV. As is demonstrated in the chart in the Appendix, what is predicted theoretically **does** manifest itself in the system of derivations of the Hungarian language; in this way we could work out a system revealing several hidden connections in an intricate area of grammatical description, which is a good way of verifying the explanatory adequacy of the theoretical framework.

As for further research in the area, beyond the "quantitative" problem of studying the derivational system of other languages, we would like to extend the theory to the area of **aspect** and **event(uality) structure** (Alberti–Ohnmacht 2005): we would like to base the calculation of eventuality structures of predicators upon the semantic properties of derivational operations producing these predicators step by step.

References

Alberti, Gábor 1988a. Tau modell (a természetes nyelv egy függvényszerkezetes modellje) [Model Tau: A model of natural language based on functions] (Műhelymunkák 4). MTA Nyelvtudományi Intézete, Budapest. 3–39.

Alberti, Gábor 1988b. Tau modell (a természetes nyelv egy sztochasztikus modellje) [Model Tau: A stochastic model of natural language]. Diploma work. Eötvös Loránd University (ELTE), Budapest.

Alberti, Gábor 1992–1993. Hatáslánc [Chains of influence]. In: Nyelvtudományi Közlemények 93 : 137–59.

Alberti, Gábor 1994. Model Tau: A formal theory of thematic roles. In: Zoltán Bánréti (ed.): Papers in the theory of grammar, 184–235. Research Institute for Linguistics of the Hungarian Academy of Sciences, Budapest.

Alberti, Gábor 1995. Role assignment in Hungarian possessive constructions. In: István Kenesei (ed.): Approaches to Hungarian 5 (Levels and structures), 13–28. JATEPress, Szeged.

Alberti, Gábor 1996. Passzivizálási művelet a magyarban [Passivization in Hungarian]. In: Néprajz és Nyelvtudomány 2:7–46.

Alberti, Gábor 1997. Argument selection. Peter Lang, Frankfurt am Main.

Alberti, Gábor 1998a. On passivization in Hungarian. In: Casper de Groot – István Kenesei (eds): Approaches to Hungarian 6 (Papers from the Amsterdam conference), 103–22. JATEPress, Szeged.

Alberti, Gábor 1998b. Struktúrák a vonzatszerkezetben [Structures in argument structure]. In: László Büky – Márta Maleczki (eds): A mai magyar nyelv leírásának újabb módszerei III [Recent methods in the description of contemporary Hungarian 3], 185–221. JATEPress, Szeged.

Alberti, Gábor 2000. Jelentésleírás — „életfogytiglani" DRS ben [Describing meanings — in "lifelong" DRSs]. In: Tamás Gecső (ed.): Lexikális jelentés – aktuális jelentés, 12–25. Tinta Kiadó, Budapest.

Alberti, Gábor – Magdolna Ohnmacht 2005. Aspect and eventuality structure in Hungarian. Paper presented at the Seventh International Conference on the Structure of Hungarian, Veszprém, 29–31 May 2005.

Bene, Annamária 2005. Az igék bennható–mediális–tranzitív felosztásának alkalmazhatósága magyar szintaktikai és morfológiai sajátosságok magyarázatában [Applicability of the classification of verbs into unergative, unaccusative and transitive groups to explaining Hungarian syntactic and morphological phenomena]. Doctoral dissertation, Eötvös Loránd University (ELTE), Budapest.

Bresnan, Joan – Jonni M. Kanerva 1989. Locative inversion in Chichewa: A case study of factorization in grammar. In: Linguistic Inquiry 20:1–50.

Dowty, David R. 1991. Thematic proto-roles and argument selection. In: Language 67:547–619.

Fillmore, Charles J. 1968. The case for case. In: Emmon Bach – Robert T. Harms (eds): Universals in linguistic theory, 1–88. Holt, Rinehart and Winston, New York.

Grimshaw, Jane 1990. Argument structure (Linguistic Inquiry Monograph 18). MIT Press, Cambridge MA.

Gruber, Jeff 1965. Studies in lexical relations. Doctoral dissertation, MIT, Cambridge MA. Published as Lexical Structures in Syntax and Semantics. North Holland, Amsterdam, 1976.

Jackendoff, Ray 1987. The status of thematic relations in linguistic theory. In: Linguistic Inquiry 18:369–411.

Kenesei, István 2000. Szavak, szófajok, toldalékok [Words, parts of speech, affixes]. In: Kiefer (2000, 75–136).

Kiefer, Ferenc (ed.) 2000. Strukturális magyar nyelvtan 3. Morfológia [A structural grammar of Hungarian, Vol. 3: Morphology]. Akadémiai Kiadó, Budapest.

Kiefer, Ferenc – Mária Ladányi 2000a. A szóképzés [Word formation]. In: Kiefer (2000, 137–64).

Kiefer, Ferenc – Mária Ladányi 2000b. Morfoszintaktikailag semleges képzések [Morphosyntax-neutral derivation]. In: Kiefer (2000, 165–214).

Komlósy, András 1982. Deep structure cases reinterpreted. In: Ferenc Kiefer (ed.): Hungarian general linguistics, linguistic and literary studies in Eastern Europe 4, 351–85. John Benjamins, Amsterdam & Philadelphia.

Komlósy, András 1992. Régensek és vonzatok [Valence and government]. In: Ferenc Kiefer (ed.): Strukturális magyar nyelvtan 1. Mondattan [A structural grammar of Hungarian, Vol. 1: Syntax], 299–527. Akadémiai Kiadó, Budapest.

Komlósy, András 2000. A műveltetés [Causativization]. In: Kiefer (2000, 215–92).

Laczkó, Tibor 2000a. A melléknévi és határozói igenévképzők [Forming adjectival and adverbial participles]. In: Kiefer (2000, 409–52).

Laczkó, Tibor 2000b. Az ige argumentumszerkezetét megőrző főnévképzés [Nominal derivation retaining the argument structure of the base verb]. In: Kiefer (2000, 293–407).

Laczkó, Tibor 2000c. Zárójelezési paradoxonok [Bracketing paradoxes]. In: Kiefer (2000, 619–52).

Parsons, Terence 1995. Thematic relations and arguments. In: Linguistic Inquiry 26: 635–62.

Perlmutter, David 1978. The unaccusative hypothesis and multiattachment: Italian evidence. Paper presented to the Harvard Linguistics Circle, May 9, 1978.

Rappaport, Malka – Beth Levin 1988. What to do with theta-roles? (Syntax and Semantics 21: Thematic Relations). Academic Press, New York.

Szabolcsi, Anna 1992. A birtokos szerkezet és az egzisztenciális mondat [The possessive construction and the existential sentence]. Akadémiai Kiadó, Budapest.

Tenny, Carol L. 1994. Aspectual roles and the syntax–semantics interface. (Studies in Linguistics and Philosophy 52). Kluwer, Dordrecht.

Tóth, Ildikó 2000. Inflected infinitives in Hungarian. Doctoral dissertation, Tilburg University.

Williams, Edwin 1995. Theta theory. In: Gert Webelhuth (ed.): Government and binding theory and the minimalist program. Principles and parameters in syntactic theory, 99–124. Blackwell, Cambridge MA & Oxford.

Zsilka, János 1966. A magyar mondatformák rendszere és az esetrendszer [The system of Hungarian sentence structures and cases] (Nyelvtudományi Értekezések 53). Akadémiai Kiadó, Budapest.

Zsilka, János 1982. *De constructione* (Történet és állapot egysége a nyelvben) [*De constructione* (The unity of diachrony and synchrony in language)]. Akadémiai Kiadó, Budapest.

Appendix

The system of Hungarian (semi-) productive derivational suffixes on the basis of transition patterns of case frames (see comments in sections **3–5**)

		⟨NOM/POSS/Ø, ACC/Ø⟩	⟨NOM/Ø⟩	⟨NOM/Ø, POSS⟩
	⇗	V / [Participle]	A / Adv	N
−N −C +C +N ↓ ↓ −N −C +C +N	V	*-gAt: kavargat, ♣ borozgat (-hAt)* PRESENT PARTICIPLE: [*-Ó*, also ♣] [°newspaper dial. *-(Vt)t*, also ♣] [†*-(Vt)tA*] INFINITIVE: [*-ni: P-nek megbántania M-t / ♣ káromkodnia*] ADVERBIAL PARTICIPLE: [*-vA*$_{active}$: *látva a helyzetet / ♣ hazaérve*]	*(-ékeny/ékony)*	*-ÓjA: Mari Péter(-nek a) szeretője / elcsábítója*
	A	*-Vdik: sötétedik* *-Ul: kizöldül* *-(Vl)lik: sárgállik*	*-(A)s: zöldes* *-AtlAn, -tAlAn: bátortalan,szerénytelen* *-An: részegen* *-Ul: cudarul*	conversion: *angol*$_N$ *(szerető*$_N$*, autószerelő*$_N$*)*
	N	*-(Vs)kodik: basáskodik, katonáskodik*	(conversion: *király*$_A$)	*-(V)(cs)kA* and other diminutive der. suf.
−N −C +C +N ↙ (↘) −N −C +C +N	V	*ken / fakaszt* *♣ (be)épül, zsong / fakad* *♣ megtelefonál* "pseudo-objects"		*♣ -Ó*$_3$*: daráló* (place)
	A			
	N			♣ possessive: *-(j)A*?
−N −C +C +N ↘ (↙) −N −C +C +N	V	caus.: *-(t)At: ásat* *♣ dolgoztat vkit (vkivel/vhol)* (non-prod. *elsüllyeszt*)		*-Ás: az o.*$_{PT}$ *simogatása* *♣ -Ás: az o.*$_{AG}$ *sim-a*
	A	*-ít: drágít* *(-Vll: drágáll)*		*-sÁg: szépsége*
	N		*-i: pécsi*	*-sÁg: királysága, tanársága*
−N −C +C +N ↘ ↘ −N −C +C +N	V	intransitivization: *eszik, olvas* refl.: *fésülködik*	*-Ó: a szerető fiad*	*-Ó*$_{1/2}$*: daráló* (worker / instrument)
		(patient) incorp.: *autót szerel,* *♣ fiú születik / vendég érkezik*	*-Ó: autószerelő*$_A$ *(brigád)*	
	A			
	N			possessor incorp.: *Ady-vers, autószerelés*
−N −C +C +N ↙ −N −C +C +N	V	"middle": *elkenődik / megoldódik* †*-(t)Atik: riporter keresztetik* PAST PARTICIPLE: [*-(V)(t)t: elgörbített / ♣ elgörbült*] FUTURE PARTICIPLE: [*-AndÓ*] ADV. PRT.: [*-vA*$_{állapot}$: *be van festve, ♣ be van rúgva – *oda van futva*]	*-hAtÓ: megoldható* *-hAtAtlAn: megoldhatatlan,* *(♣ kiapadhatatlan – *kimerülhetetlen *elmehetetlen)* *-AtlAn: megoldatlan*	*(-mÁny: rakomány, ♣ fejlemény)*
		⟨♣ German passivization⟩ ADVERBIAL PARTICIPLE: [°*-vA*$_{hasch.}$: *♣ ki van takarítva, be van ágyazva*]		
	A			
	N	*-l: kapál* *-z: gereblyéz*	*-Vs: gond-* *-(A)tlAn, -t(A)lAn* *-(j)Ú: hosszú hajú*	

Acta Linguistica Hungarica, Vol. 53 (1), pp. 41–51 (2006)
DOI: 10.1556/ALing.53.2006.1.2

ON THE HISTORICAL BACKGROUND OF HABITIVE AND IZAFET CONSTRUCTIONS IN HUNGARIAN*

LÁSZLÓ HONTI

Dipartimento di Glottologia e Filologia Classica
Università degli Studi di Udine
Via Mazzini 3
I–33100 Udine
Italia
litnoh@uniud.it

Abstract: This paper deals with two ways of expressing possessive relationships, their morphological make-up and the possible circumstances of their emergence. One of these is the habitive construction ('X has Y'), whereas the other is the attributive possessive construction ('X's Y, the Y of X'). The former is a clause, whereas the latter is a phrase. It will be argued that both types of constructions may have emerged in the Uralic languages without the contribution of any foreign influence, but as far as the retention of the latter is concerned, foreign influence may have had a role in it in Uralic languages that were engaged in intensive Uralic–Turkic linguistic contacts.

Keywords: possessive, izafet, habitive, language contact, case marking

In the Uralic languages, two types of constructions serve as a clause-level expression of a possessive relationship:

(a) the possessed subject is in the nominative, while the nominal standing for the possessor has some case suffix that may be either grammatical (nominative, genitive, or dative) or else adverbial (lative, locative, or ablative); the possessed noun may or may not bear a personal possessive suffix, and the clause may or may not involve a copula; or

(b) the possessed noun is the direct object of a verb meaning 'have'.

* A preliminary version of this paper was presented at the Seventh International Congress of Hungarian Linguistics, Budapest, August 29–31, 2004.

1216–8076/$ 20.00

These main types comprise several subtypes. Neither of these types is exclusively used in the individual Uralic languages, although one of them is usually the dominant one. The habitives of certain Finno-Ugric languages have been claimed to be related to similar constructions occurring in Indo-European languages spoken in adjoining territories.

Phrase-level possessive relationship, on the other hand, is expressed by an attributive construction in which the possessive attribute is a nominal in some grammatical or adverbial case (nominative or genitive, respectively localis), while the possessed nominal may or may not have a possessive personal suffix. The subtype in which the possessor is not a personal pronoun and the possessed noun bears a personal suffix has been referred to by several researchers, on the basis of the name of a similar construction in Turkic languages, by the technical term "izafet". With respect to izafet constructions in some Uralic languages, some researchers suggested the possible influence of Turkic languages that are (or used to be) spoken in neighbouring areas.

1. One of the characteristic features of Uralic languages is that they tend not to make use of a verb meaning 'have', although such a verb is not completely unattested in them, either (it is rather generally used in Ob-Ugric).

1.1. The habitive construction most often attested in various Uralic languages is of the type exemplified by Latin *mihi liber est* 'I have a book' [lit. to-me book is]. That is, it involves the verb 'be', a nominal—mostly case-marked or followed by a postposition—standing for the possessor, and a possessed nominal that often has a possessive personal suffix attached to it. The individual languages may mark the possessor in diverse ways. In Finnic, Mordvin, Cheremis, and Samoyedic, reflexes of Proto-Uralic genitive **n* and/or dative **ń* are normally used in habitive constructions; Hungarian started to use dative *-nak/-nek* (of a lative postpositional origin) for that purpose; Finnic languages—except for Kurland Livonian—use the allative, the adessive, or the ablative as a secondary device, while the locative is used in Lappish, the adessive or the ablative in Permic, and a locative postposition in Selkup. In the Finnic–Lappish group, the possessed noun never takes possessive personal suffixes, whereas in Hungarian it always does. In the rest of the languages, both solutions are used, with one of them dominant, at most. The verb is usually overt, but its position is not fixed across languages and across constructions.

Thus, the scheme of the Uralic habitive construction is "Noun-Ø/Gen/Adv/Postp + Noun-Ø/Px (+ 'is')".

The morphological shapes of habitive and attributive possessive constructions are interrelated as sown below in a tabular form:

	IN HABITIVE CONSTRUCTIONS	IN ATTRIBUTIVE CONSTRUCTIONS
Balto-Finnic	case markers involving *l*, genitive	genitive
Lappish	static local cases	genitive
Mordvin	genitive	genitive
Cheremis	genitive	nominative, genitive
Permic	case markers involving *l*, nominative	case markers involving *l*, nominative
Ob-Ugric	nominative, locative, postposition	nominative, locative
Hungarian	dative (> genitive)	nominative, dative (> genitive)
Northern Samoyedic	nominative, genitive	genitive
Southern Samoyedic	nominative, genitive, postposition	genitive

Note: The dative *and* genitive interpretation of Hungarian *-nak/-nek* is not uncontroversial in the literature (cf. e.g., Korompay 1991, 301–2 vs. Kiefer 2000, 577–8).

In addition to a verbal predicate and the subject (that is, the possessed noun), habitive constructions usually involve the possessor, too. Whenever the possessed noun bears a possessive personal suffix—which is commonly the case outside Balto-Finnic and Lappish—, pronominal possessors are represented by personal suffixes, hence the possessive personal pronouns need not be overt in such cases. Overt possessors are in the genitive (going back to dative) in Finnic; also in Finnic as well as in Lappish it may be the case that the possessor is in the localis and the possessed noun has no personal suffix (within this group, Kurland Livonian and Southern Lappish are exceptions in that the possessor is suffixed by *-n* in the former and by the genitive suffix in the latter); in Volgaic and Permic languages, the possessed noun does not necessarily have a personal suffix but then the possessor has to be case-marked; Ob-Ugric languages, lacking the genitive, use the nominative instead.

Examples: Finnish *minu-n on nälkä-Ø* ~ *minu-lla on nälkä-Ø* 'I am hungry', *minu-lla on lapsi-Ø* 'I have a child', Livonian K *izā-n um tidār-Ø*

'the man has a daughter', Lappish N *makkár beatnagat-Ø du-s leat?* 'what kind of dogs do you have?', *mu-s dat gal leat buoret beatnagat-Ø* 'I do have good dogs', *mu-st læ manna-Ø* 'I have a child', Lappish S *muw*[gen] *mana-Ø* 'I have a child', Mordvin *eŕźa-ń uĺi suk-azo* 'an Erzan has a dog', Cheremis *[nun-ân] ikšə̑wə̑-št uke* 'they have no child', *məń-ən u kńiyä-Ø ulə̑* 'I have a new book', Votyak *[men-am] nįĺ-į vań* 'I have a daughter', Zyryan *joma-lęn pi-įs abu* 'the wizard has no sons' ~ *nį-lęn em kujim pi-Ø* 'they have three sons', Ostyak Vj *mä-Ø wə̑γ-am ĕntim* ~ *mĕn-nə wăγ-Ø ĕntim* 'I have no money', Kr *χuj pĕŋətnə wăγ ut* 'the man has money' ['bei dem Mann ist Geld'], Vogul TJ *näjär̈-Ø püw-ī āləs* 'the emperor had a son', *püw-ǟn āl* 'they have a son', TČ *jükǟ pält ǟwīťī ǟls* 'the woman had a daughter' ['bei der Frau war ihre Tochter'], Hungarian *lov-am van* 'I have a horse' [horse-my is], *az apá-nak ház-a van* 'the father has a house' [the father-dat house-his is] ~ *az apá-nak nincs ház-a* 'the father has no house' [the father-dat is-not house-his].

1.2. The claim that the possessor-marking function of the dative is an ancient feature is supported by the fact that in several of the Uralic languages the dative suffix **ń* is used in that role; note that constructions involving more recent dative suffixes are also attested. E.g., Cheremis *kuruk-lan üšək ulo* 'the hill has a shadow', *kuku-lan šəδaŋəm βüδalš(-at) ulo* 'the cuckoo has a wheat sower'. In Hungarian, the situation is similar: *a hegy-nek árnyéka van* 'the hill has a shadow', *a kakukk-nak búzavetője van* 'the cuckoo has a wheat sower'. The appropriateness of the dative to express a possessive relationship is also clearly demonstrated by the closest relatives of Hungarian: Ostyak and Vogul. In these languages, the dative of personal pronouns has a role in the expression of a possessive relationship if the possessed noun is in the singular — although not in habitive or attributive possessive constructions but in a predicative role (and perhaps also in that of a complement, as in the case of the Hungarian possessive pronouns *enyém* 'mine', *tiéd* 'yours', etc.):

(a) Ostyak N *манэм* 'to me; my': *тами ма шукем!* — *антэ, манэм!* 'Это моя крошка! — Нет, моя!'.

(b) Ostyak Kaz *naŋen* 'to you; your', *łuveł* 'to him; his': *tam łaŋki muj naŋen muj łuveł* 'эта белка или твоя, или его'.

(c) Ostyak Kaz *mănem* 'to me; my', *măŋew* 'to us; our': *χŏj pŭťįĺka?* — *mănem.* — *măŋew.* 'Whose bottle [is this]? — Mine. — Ours'.

(d) In Vogul, it is likewise the dative-lative forms of personal pronouns that figure as possessive pronouns, cf. N *ānəm* 'me; to me; my', *tawe* 'him; to him; his', *mēnmēn* 'us; to us; our (du. 1)', *mānaw* 'us; to

us; our (pl. 1)': *nēpək ānəm* 'бумага моя', *pisaľ tawe* 'ружье его', *kolәγ mēnmēn* 'дома (дв.) наши (дв.)', *sālit mānaw* 'олени наши'. This construction is also illustrated by the following Southern Vogul example: TJ *äm(ǟn) waŋkā k̥irt kālәŋ̥ kǟšəm ǟ sows* 'ich hatte keine Lust, in dem Brunnen zu sterben', 'mir kam [eigtl. wurde] nicht meine Lust... zu sterben'. The sentence can be interpreted in two ways, depending on whether it includes a nominative or a dative pronoun: 'I had no desire to...' and 'there was no desire in me to...'.

1.3. As the foregoing suggest, in habitive constructions of present-day Uralic languages, the possessor often bears a case suffix (originally localis), not infrequently that of the dative; dative and lative are also known in that role in some Indo-European languages, among others (similarly with respect to attributive possessive constructions). The Indo-European languages, or their proto-languages, originally lacked a verb for 'have'; rather, they formed their habitive constructions with the dative of the possessor noun and the verb for 'be' (see Gamkrelidze–Ivanov 1984, 288–9, and references therein). In Latin, if the possessed noun is emphasised, the verb is some finite form of 'be' and the possessor is in the dative (dativus possessivus), e.g., *mihi liber est* 'I have a book'; but if the possessor is made prominent, the genitive (genitivus possessivus) is used: *patris est domus* 'the house belongs to the father'. According to Havers (1911, 319 and footnote), this is "dativus sympatheticus" that occurs both in habitive and in attributive possessive constructions in some Indo-European languages; in Sanskrit and Old Greek, it is restricted to pronouns, especially personal pronouns, but in Vulgar Latin, in Germanic, and in Balto-Slavic, it was extended to nouns, and in Romanian, Albanian, Bulgarian, and Modern Greek, the genitive and the dative became syncretised (as in the Finnic and Volgaic groups of Uralic languages). E.g., German dialectal *er ist ein Vetter zu mir, wem sein Hut* 'wessen Hut', German *dem Vater sein Sohn* ~ *des Vaters sein Sohn*, Dutch *de vader zijn zoon* 'id.', English *server to his master*, Old English *him on þet heafod* 'to-him on the head (= on his head)', French *ce livre est à moi, le cheval à mon père*, Italian *descriver fondo a tutto l'universo* 'to describe the essence of the whole world', Russian *ему шесть лет, доктор ему осмотрел горло*, Latvian *kas tev ir?* 'who/what do you have?'. Thus, the dativus possessivus found in Finno-Ugric languages is not unprecedented.

1.4. With respect to the origin of habitive and attributive possessive constructions involving dative or locative prepositional/postpositional

phrases, radically diverse views have been expressed in terms of Finno-Ugric–Indo-European contacts, listed here without comment:

(a) In the background of the Latvian and Russian constructions, some people claim there must be a Finno-Ugric substratum or adstratum, whereas others deny this; indeed, the question arose of whether the Finnic construction (Finnish *minulla on...*) is not of Russian origin to begin with.

(b) With respect to the function of the Livonian dative, perhaps some Latvian influence could be considered.

(c) Given that, in Ostyak, the locative of personal pronouns can be used to express the possessor, this probably shows the influence of Russian.

(d) A similar view has been expressed with respect to the Vogul postpositional construction.

(e) The Selkup postpositional phrase may be due to Russian influence, too; in this case, Selkup may have extended the earlier construction 'my reindeer is' by an item corresponding to Russian *у меня* but retained the possessive suffix as well.

(f) According to Nikolaeva (2002, 283–4), Hungarian dativus possessivus is the result of convergence with what she calls "European" languages; but in the same paper, she also claims that Hungarian "copied" that construction from neighbouring languages.

As far as I know, the origin of Hungarian dativus possessivus has been satisfactorily explained by historical linguists and, mutatis mutandis, this applies to events that may have taken place in long-bygone periods. Therefore, no obscure links are to be suspected with respect to the origin of the corresponding Uralic and Indo-European constructions.

The nominal bearing a locative suffix or a locative postposition may originally have served to express the whereabouts of the denotatum of the possessed noun rather than the identity of the possessor, as can still be felt e.g., in Finnish *meillä on vieraita* 'we have guests' ~ 'there are guests in our place'; this suggests that such constructions may have arisen in a natural, spontaneous manner.

2. Turning to **izafet constructions**, I have to explain very briefly what that term is meant to express since the notion is not widely known, as far as I am aware (cf. Yartseva 1990, 172). Given that the term is used for Uralic languages with reference to patterns in Turkic languages, I considered primarily definitions that interpret it in the framework of discussions of

Turkic languages. The scanty Turkological literature on the issue that I had access to, as well as references within Uralistics, suggest that an attributive possessive construction counts as an instance of izafet if the possessive relationship is marked (at least) on the possessed noun, and its marker is a possessive personal suffix, e.g., Hungarian *a ház tete-je* [the house roof-its], *a ház-nak a tete-je* [the house-dat the roof-its] 'the roof of the house'. Hence, izafet is one of the possessive constructions, of which Uralic languages exhibit at least four. These are as follows:

(a) The possessive relationship is not morphologically marked on either part of the construction (this, then, is not izafet); often, compounds and compound-like phrases belong here, e.g., Hungarian *háztető* 'roof of house' [lit. house-roof]; Mordvin *utom keŋkš* 'дверь амбара (букв.: амбар-дверь)'; Cheremis *pünžö wujə-što* 'on top of the hill'; Votyak *škola sad* 'школьный сад'; Zyryan *pįzan dęra* 'скатерть; сальфетка'; Ostyak Vj *ni poťiŋka* 'the woman's shoes'; Vogul *puŋk-sow* 'scalp'; Selkup *pōt pīry* 'высота дерева', that is: "Noun-*Ø* + Noun-*Ø*".

(b) A possessive morpheme only occurs on the attributive constituent (this is not izafet either, since the head noun does not exhibit any mark of possessivity), e.g., Finnish *talo-n katto* 'roof of the house'; Cheremis *kuγiž-än əδər-Ø* 'the king's daughter'; Ostyak Tra *ťårəs-nə jěŋk-Ø-a* 'into (the water of) the sea'; Yurak-Samoyed *ṕa-ʔ hadeʔ-Ø* 'смола дерева'; Yenisey Samoyed *abā-ń pag-ę* 'паница моей старшей сестры', that is: "Noun-Cx + Noun-*Ø*".

(c) The possessive relationship is only marked on the possessed noun, by a possessive personal suffix (this is izafet, since the morpheme expressing the relationship between the constituents is located on the head noun), e.g., Hungarian *a ház-Ø tete-je* 'roof of the house'; Ostyak Vj *räť-Ø* ɔγ-*ə̂l* 'head of the old man'; Vogul N *lū-Ø puŋk-e* 'head of the horse'; Zyryan *kaťę-Ø ź́ugj-as* 'in Kätchens Schlinge', that is: "Noun-*Ø* + Noun-Px".

(d) Both the attribute and the head noun exhibit morphemes expressing the possessive relationship (this is also izafet, since one of the morphemes expressing the relationship between the constituents is located on the head noun), e.g., Hungarian *a ház-nak a tete-je* 'the roof of the house'; Cheremis *kuγižä-n əδər-žə* 'the king's daughter'; Votyak *kolhoz-len muzjem-ez* 'земля колхозная'; Ostyak DN *ťăras-nat wăχ-ət* 'das Geld des Kaufmannes'; Yurak-Samoyed *ńa-kan męti-da* 'упряжные олени моего старшего брата (букв. брата-моего упряжные-его)'; Yenisey Samoyed *kɛδɛr-' koba-δa* 'шкура дикого оленя', that is: "Noun-Cx + Noun-Px".

2.1. Bereczki (1983a, 65; 1984, 308), the author who has discussed the matter the most thoroughly, mentions Volgaic and Permic languages, as well as Hungarian, as languages that employ izafet constructions, noting that—except for Hungarian—these languages also exhibit the original Finno-Ugric construction involving no possessive personal suffixes. Simonyi (1914, 139) was furthermore aware that the same construction occurred in Vogul and in Ostyak, too. Thereby the insight of Beke (1914–1915, 21) the author who first investigated this issue, saying that izafet is known to exist (albeit not exclusively) in all Finno-Ugric languages except for Finnic and Lappish, appears to have proved correct.

2.2. For curiosity's sake let me add here that (of the Indo-European languages spoken in Europe) possessive phrases reminiscent of izafet constructions of Uralic and Altaic languages are not totally unfamiliar in Germanic and in Romance, either. The literature I have looked at knows of only two such groups of Indo-European languages, e.g., German *wem sein Hut* 'wessen Hut', *dem Vater sein Haus* ~ *des Vaters sein Haus* ~ Dutch *de vader zijn huis*, Italian *della mia sopravvesta il suo colore* 'the colour of my coat'. In Germanic, the phrase-initial possessor is topicalised (Ramat 1986, 587), and the possessor is (at least historically speaking) in the dative, whereas in Italian the genitive form (*della*) is used; the possessed noun is determined by a third-person possessive attributive pronoun in all three languages cited. That is, these phrases exactly match the izafet constructions of Uralic languages, with the difference that they refer back to the possessor by a possessive pronoun rather than by a possessive personal suffix.

2.3. The emergence of the izafet construction does not have a large literature but it includes at least three different views. Some authors claim that it is probably of Proto-Uralic origin (Hajdú 1987, 222–3, cf. Benkő 1979, 57; 1988, 24–5), others trace it back to contacts with Turkic languages, at least with respect to Volgaic and Permic languages, and Hungarian (Beke 1914–1915, 21–7; Bereczki 1983b, 214; Rédei 1980, 86), whereas Fokos (1939, 16) considers it to be a result of internal development. With respect to Samoyedic, I am unaware of attempted explanations of origin.

Ugric languages may have introduced the izafet construction due to the fact that the genitive they had inherited from Proto-Finno-Ugric had been lost and in some cases they needed to reflect the grammatical relationship in their possessive constructions, or to make the possessor more prominent by topicalising it, but it was not possible to unambigu-

ously indicate by a case marker, due to lack of genitive, that two nouns formed a possessive construction. In such cases, Ugric languages could (and can) disambiguate the syntactic relationship between the two constituents by a third-person possessive suffix attached to the possessed noun, topicalising the possessor; the use of the person marker was generalised in Hungarian e.g., *(a) fiú láb-a* 'the foot of the boy', whereas in Ob-Ugric it remained an occasional topicalising device. This may have been reinforced by the analogy of constructions involving a pronominal possessor; as Fokos (*ibid.*) pointed out, "constructions like *apám háza* 'my father's house' may have been modelled after constructions like *ő háza* 'his house'." Beke (1914–1915, 21–7) explained the emergence of izafet constructions in these Finno-Ugric (Volgaic, Permic, Ugric) languages by an influence of Volga Turkic languages, and Bereczki (1983a, 65) did the same with respect to Permic and Volgaic languages, whereas for Hungarian he thought it was possible that "it was a result of internal development." The possessive personal suffix also emerged as an alternative in Permic and Volgaic languages. Given that its overt occurrence is obligatory in Hungarian, we can say that this language is more consistent in this kind of possessor marking than Turkic languages are. That is, Hungarian is the ideal izafet language. Another point in favour of the spontaneous development account is that not all Uralic languages have been spoken next to Turkic languages for a thousand years or so (in the case of Hungarian, such coexistence was but a moment's episode, in a historical perspective). I think, in accordance with Fokos, Hajdú, and Benkő, that the emergence and/or retention of izafetism was underpinned by internal necessity.

An argument supporting the Turkic origin of izafet constructions in some Finno-Ugric languages is that they are unknown in Finnic and Lappish that were not in contact with Turkic languages. Thus, provided I do not accept the alleged role of Turkic influence in this respect, I have to find an answer to the question of why this construction is not attested in the westernmost languages of the family. In Finnic or Lappish, topicalisation of this kind did not occur, although it undoubtedly might have, rather—at least in Finnish—it was the possessor with a genitive suffix that acquired additional prominence. This can be explained, I think, by the fact that the genitive suffix had been retained, a necessary but not sufficient condition for izafet constructions not to occur in languages of the westernmost group. That construction is not quite generally used in other Uralic languages, either, with the sole exception of Hungarian. But

then, this is the single Uralic language that had the shortest contact of all with Turkic languages. Whether the izafet construction is of a Proto-Uralic origin or it emerged in the separate lives of individual languages or language groups, I would be reluctant to claim that its emergence in just these languages is due to mere chance. Indeed, intensive contacts with Turkic languages may have stimulated either the retention of original izafet constructions or the reinforcement of a trend of development triggered by an internal need. Thus, I do not deny the possible role of Uralic–Turkic contacts or bilingualism, yet I do not at all see it as a decisive factor. In that respect, it is also worth noting that izafet(-like) constructions can be found in European Indo-European languages that had no Turkic contacts whatsoever. In sum: the izafet construction of Uralic languages may be a result of internal changes, rather than that of an adopted foreign model.

3. Both habitive and izafet constructions may have emerged in the Uralic languages without the contribution of any foreign influence, but as far as the retention of izafet is concerned, foreign influence may have had a role in it in Uralic languages that were engaged in intensive Uralic–Turkic linguistic contacts.

References

Beke, Ödön 1914–1915. Türkische Einflüsse in der Syntax finnisch-ugrischer Sprachen. In: Keleti Szemle 15 : 1–77.

Benkő, Loránd 1979. Über die Bezeichnung des ungarischen Possessivattributs. In: Christoph Gläser – János Pusztay (eds): Festschrift für Wolfgang Schlachter zum 70. Geburtstag. Veröffentlichungen der Societas Uralo-Altaica. Band 12, 57–64. Otto Harrassowitz, Wiesbaden.

Benkő, Loránd 1988. A magyar birtokos jelzős szerkezet jelölésének történetéből [From the history of Hungarian attributive possessive constructions]. In: Magyar Nyelv 84 : 24–31.

Bereczki, Gábor 1983a. A török nyelvek hatása a magyarra [The influence of Turkic languages on Hungarian]. In: Mémoires de la Société Finno-ougrienne 185 : 59–72.

Bereczki, Gábor 1983b. A Volga–Káma-vidék nyelveinek areális kapcsolatai [Areal contacts of languages of the Volga–Kama region]. In: János Balázs (ed.): Areális nyelvészeti tanulmányok [Studies in areal linguistics], 207–36. Tankönyvkiadó, Budapest.

Bereczki, Gábor 1984. Die Beziehungen zwischen den finnougrischen und türkischen Sprachen im Wolga-Kama-Gebiet. In: Nyelvtudományi Közlemények 86 : 307–14.

Fokos, Dávid 1939. Finnugor–török mondattani egyezések [Finno-Ugric–Turkic syntactic correspondences]. In: Magyar Nyelvőr 68 : 13–9.

Gamkrelidze, T. V. – V. V. Ivanov [Гамкрелидзе, Т. В.-В. В. Иванов] 1984. Индоевропейский язык и индоевропейцы. Реконструкция и историко-типологический анализ праязыка и протокультуры. Издательство Тбилисского университета, Тбилиси.

Hajdú, Péter 1987. Die uralischen Sprachen. In: Péter Hajdú – Péter Domokos: Die uralischen Sprachen und Literaturen, 21–450. Budapest & Hamburg, Akadémiai Kiadó & Helmut Buske Verlag.

Havers, Wilhelm 1911. Untersuchungen zur Kasussyntax der indogermanischen Sprachen. Untersuchungen zur indogermanischen Sprach- und Kulturwissenschaft 3. Verlag Karl J. Trüber, Straßburg.

Kiefer, Ferenc 2000. A ragozás [Inflection]. In: Ferenc Kiefer (ed.): Strukturális magyar nyelvtan 3. Morfológia [A structural grammar of Hungarian, Vol. 3: Morphology], 569–618. Akadémiai Kiadó, Budapest.

Korompay, Klára 1991. A névszóragozás [Declension]. In: Loránd Benkő – Erzsébet E. Abaffy – Endre Rácz (eds): A magyar nyelv történeti nyelvtana. 1. kötet: A korai ómagyar kor és előzményei [A historical grammar of Hungarian. Volume 1: Early Old Hungarian and its antecedents], 284–318. Akadémiai Kiadó, Budapest.

Nikolaeva, Irina 2002. The Hungarian external possessor in a European perspective. In: Rogier Blokland – Cornelius Hasselblatt (eds): Finno-Ugrians and Indo-Europeans: Linguistic and literary contacts (Studia Fenno-Ugrica Groningana 2), 272–285. Shaker Publishing, Maastricht.

Ramat, Paolo 1986. The Germanic possessive type *dem Vater sein Haus*. In: Dieter Kastovsky – Aleksander Szwedek (eds): Linguistics across historical and geographical boundaries. In honour of Jacek Fisiak on the occasion of his fiftieth birthday. Volume 1. Linguistic theory and historical linguistics. Trends in linguistics. (Studies and Monographs 32), 579–90. Mouton de Gruyter, Berlin & New York.

Rédei, Károly 1980. [Diskussionsbeitrag]. In: Gyula Ortutay (ed.): Congressus Quartus Internationalis Fenno-Ugristarum Budapestini habitus anno 1975. Pars II, 85–7. Akadémiai Kiadó, Budapest.

Simonyi, Zsigmond 1914. A jelzők mondattana [The syntax of attributes]. Magyar Tudományos Akadémia, Budapest.

Yartseva, V. N. (ed.) [Ярцева, В. Н. (отв. ред.)] 1990. Лингвистический энциклопедический словарь. Советская энциклопедия, Москва.

Acta Linguistica Hungarica, Vol. 53 (1), pp. 53–76 (2006)
DOI: 10.1556/ALing.53.2006.1.3

THE STRUCTURE OF SELF-REPAIRS IN THE SPEECH OF HUNGARIAN LEARNERS OF ENGLISH

JUDIT KORMOS

Department of English Applied Linguistics
Eötvös Loránd University
19–21 Ajtósi Dürer sor
H–1146 Budapest
Hungary
kormos.j@chello.hu

Abstract: The study reported in the paper investigates the structure of L2 self-corrections in the speech of 30 Hungarian learners of English and 10 Hungarian native speakers. The aim of the research is to examine what the well-formedness of the corrections, the use of editing terms, the placement of cut-off points and the effect of the participants' level of proficiency on the structure of self-repairs reveal about the psycholinguistic processes of speech production. The results of the study lend additional support for modular models of speech production (e.g., Levelt 1983, 1989; Levelt et al. 1999) and reveal an important role of pragmatic constraints in psycholinguistic processing.

Keywords: speech production, self-correction, monitoring, editing, second language

1. Introduction

The self-correction behaviour of second language learners has already been explored from a number of aspects. In earlier studies it has been argued that the investigation of self-repairs in L2 speech can yield insights into mechanisms of speech production (e.g., Kormos 1999, 2000a,b; Levelt 1983, 1989; Poulisse 1999; Poulisse–Bongaerts 1994; van Hest 1996), processes of speech automatization (Kormos 2000b) and the allocation of attention (Kormos 2000a). In this regard, the distribution of self-repairs and its relation to the frequency of errors and to the development of proficiency as well as the timing of self-corrections was studied

by a number of researchers (for a comprehensive review of self-repair research see Kormos 1999). A less frequently investigated aspect of self-corrections in L2 is the structure of self-repairs. The existing studies in this field mainly used either native speakers of Dutch or Dutch learners of English and were descriptive in nature (with the exception of Brédart's (1991) study, whose participants were French). No inferential statistical analyses to support the claims were carried out in earlier research, and the effect of L2 proficiency on the structure of self-corrections was not investigated either. The present research aimed to fill in this gap and used Hungarian native speakers as well as Hungarian learners of English at three different levels of proficiency as participants (for an overview of Hungarian speakers' self-repair behaviour see Gósy 2004; Huszár 2005). The objectives of the research were to test whether the hypotheses set up in earlier research concerning native speakers of an Indo-European language are also tenable for native speakers of Hungarian as learners of English. The study also intended to investigate whether English language competence has any marked effect on the structure of self-corrections. The results of the study can help further refine or support existing theories of speech production and can shed new light on certain pragmatic aspects of self-corrections.

2. Review of literature

2.1. A brief overview of theories of speech production

Speech production researchers all agree that language production has four important components: (1) **conceptualization**, that is, planning what one wants to say, (2) **formulation**, which includes the grammatical, lexical and phonological encoding of the message, (3) **articulation**, in other words, the production of speech sounds and (4) **self-monitoring**, which involves checking the correctness and appropriateness of the produced output. There is also agreement on the questions that conceptualization, formulation and articulation follow each other in this order, and that in L1 production planning the message requires attention, while formulation and articulation are automatic, and therefore processing mechanisms can work in parallel, which makes L1 speech generally smooth and fast. Researchers also share the view that one of the basic mechanisms involved in producing speech is **activation spreading**. Activation spreading is a metaphor adapted from brain research, which is based on the finding of

neurological studies that neural networks consist of interconnected cells (neurons) that exchange simple signals called activations via the connections they have with each other (Hebb 1949). The speech processing system is assumed to consist of hierarchical levels (conceptualization, formulation, articulation), among which information is transmitted in terms of activation spreading, and of knowledge stores such as the lexicon and conceptual memory store, within which activation can also spread from one item to related items. Decisions are made on the basis of the activation levels of the so-called **nodes** that represent various units such as concepts, word forms, phonemes etc.

There exist two major theories of L1 speech production: **spreading activation** (the name is somewhat misleading because as just mentioned, both models assume that the way information is transmitted in the speech processing system is activation spreading) (e.g., Dell 1986; Dell–O'Seaghda 1991; Stemberger 1985) and **modular theories** (e.g., Fry 1969; Garrett 1976; Laver 1980; Levelt 1989; Levelt et al. 1999; Nooteboom 1980), and there are two major differences between them. The first main difference between these two theories is whether they allow for feedback between the various levels of encoding. Spreading activation models allow for the backward flow of activation from a subordinate level to the superordinate level, while in modular theories activation can only spread forwards. This means that in spreading activation theory, if an error occurs in one specific process, a warning signal is immediately issued, and activation flows upwards to the superordinate level. Processing starts again from this superordinate level. In modular models the error is not noticed at the level it is made, but only once the erroneous fragment of speech has been phonologically encoded or later when it is articulated. Therefore, in this view, bits of message that contain an error need to be encoded again from the level of conceptualization. Researchers working with modular theories argue that the processing components in the speech production system are autonomous, that is, have their own characteristic input, and they process this input independently of other components. Hence the name 'modular theory of speech production'. The second major difference between these theories concerns syntactic and phonological encoding. In spreading activation theories it is assumed that speakers first construct frames for sentences and for phonetic representations and then select the appropriate words or phonetic features for the slots in the frame. Modular models are **lexically driven**, which means that words activate syntactic building procedures, and they postulate that lexical en-

coding precedes syntactic encoding and that phonological encoding can only start once lexico-syntactic processes are ready.

2.2. Overview of earlier studies on the structure of self-repairs

The structure of self-repairs can be characterised by three variables. The first of these concerns the syntactic relationship of the reparandum and the reparatum, which provides us with information what rules govern how these two constituents of the self-correction are conjoined. The second related characteristic is the use of editing terms, which reveals by what lexical means this conjunction is expressed. Finally, the point where the erroneous or inappropriate utterance is interrupted also lends insight into the structural aspects of the repair.

Several studies have investigated whether the syntactic structure of self-repairs shows any signs of systematicity (e.g., de Smedt- Kempen 1987; Levelt 1983) and have found that the majority of self-corrections follow a specific rule, which was named the Well-formedness Rule by Levelt (1983). (Previous investigations concerning this field were discussed in detail by Kormos (1999), therefore, only a very brief summary of earlier findings will be presented here). According to the rule, "an original utterance ⟨O⟩ plus repair ⟨OR⟩ is well-formed if and only if there is a string of zero or more words ⟨C⟩ to complete the utterance so that the string ⟨OC or R⟩ is well-formed, where C is a completion of the constituent directly dominating the last element of O" (Levelt 1983, 78). In other words, this rule says that the utterance and the repair have to follow the rules of syntactic coordination. Examples (1) and (2) illustrate well-formed repairs, while examples (3) and (4) ill-formed ones.

(1) all chairs *have handles. And er sorry arms*

(2) how many people er would will *come there or here* yes?

(3) you *can this er reserve this* er er room

(4) it's almost er er the number we are *the rules er which i... which is written in our rules*

The only study that investigated the well-formedness of L2 self-repairs was conducted by van Hest (1996), who found little difference between L1 and L2 self-repairs in this respect. Following Levelt's (1983) rule of classification, 70% of the L1 self-repairs and 80% of the L2 self-repairs were

well-formed in Van Hest's corpus. On the basis of van Hest's results, one might claim that the self-repair behaviour of L2 learners also follows the well-formedness rule. Her results, however, might only be due to the fact that the syntactic structure of Dutch and that of English are similar in a number of aspects. Nevertheless, Kormos (1999) argued that van Hest's results can be taken as a proof for modular models of speech production in L2 (e.g., de Bot 1992; Levelt 1989; Levelt et al. 1999) as the findings indicate that L2 speakers re-process the whole speech plan when making a correction and do not restart the utterance from an intermediary level of production. The present research aims to substantiate this claim with data collected from Hungarian speakers.

Earlier studies on the use of editing terms concluded that the default repair structure both in L1 and L2 contains no editing term (e.g., words and expressions such as *I mean, well, or*), but if editing expressions are used, they are most likely to be filled non-lexicalised pauses (Levelt 1983; van Hest 1996). Levelt (1983) found that error-repairs were more frequently accompanied by editing terms than repairs which were concerned with the appropriacy of information. In van Hest's (1996) corpus, however, there was little difference between these two types of self-corrections in this respect. Similarly to Poulisse and Bongaerts' (1994) study, van Hest's data also suggest that L2 speakers frequently use L1 editing terms when speaking an L2 which, in most of the cases, can be regarded as unintentional code-switching. It has to be noted that these studies only used descriptive statistics and correlational analyses on the cooccurrence of editing terms and different types of self-repairs were not performed. In a recent study, Clark and Fox Tree (2002) argued that it is not only the choice of certain lexical editing terms, such as *well, I mean* etc., that is systematic, but speakers also make conscious decisions when using *uh* and *um* as delay signals. On the basis of the analysis of the spoken data in the London-Lund corpus, they claimed that *uh* signals minor delay in the speech production process, whereas *um* indicates a major delay.

The third aspect of the structure of self-corrections concerns the point of interruption of the erroneous or inappropriate utterance. The Main Interruption Rule, which was first proposed by Nooteboom (1980) and later elaborated by Levelt (1983) says that speakers stop the flow of speech immediately upon the detection of trouble. This rule suggests that regardless of the nature of trouble in the utterance, the reparandum will be interrupted immediately after the problem is perceived, and word or syllable integrity will not be respected in these cases. Levelt (1983) and

Brédart (1991), however, observed that in their study speakers tended to finish inappropriate words, while error-repairs adhered to the Main Interruption Rule. Van Hest's (1996) study with L2 learners reproduced the same result. This finding also needs to be substantiated with learners of English and speakers from a different native language background.

3. Research questions

The study reported in this paper investigated the following issues:

1. Do self-repairs in the speech of Hungarian speakers and learners of English also follow the well-formedness rule?
2. To what extent does well-formedness depend on the type of repair in the corpus?
3. How does the language proficiency of the participants affect the rate of well-formed repairs?
4. What is the relationship between the types of self-correction and the editing terms used by Hungarian speakers and learners of English?
5. What is the effect of language proficiency on the frequency and type of editing terms used by the participants?
6. Where is the point of interruption in the reparandum in the speech of Hungarian speakers and learners of English?
7. Does the language proficiency of the participants influence the placement of cut-off points?

4. Method

4.1. Settings, procedures and participants

The data for this study came from a corpus of speech samples collected from 30 Hungarian learners of English and 10 Hungarian native speakers. The database and the data collection procedures were discussed in detail in Kormos (2000a,b), therefore, the research design will be described here only briefly. Participants of the study performed a role-play task with the researcher being the interlocutor. The task was recorded and an immediate retrospective interview followed the role-play activity. The interview was conducted in accordance with the guidelines drawn up by Ericsson and Simon (1980; 1993). A C-test,[1] which had been validated

[1] In a C-test, the second half of every second word is missing.

by Dörnyei and Katona (1992) was administered to all the participants to measure their level of proficiency.

Participants of the study were all native speakers of Hungarian aged between 16 and 35. Ten students were learners in a language school and their language proficiency was at the intermediate level. Twenty participants were English majors at Eötvös Loránd University, Budapest, whose English competence was advanced. Half of the English majors performed the task in Hungarian and half of them in English. (It was not possible to collect Hungarian and English data from the same participants because of the strong effect task-repetition would have had on performance). Ten students were participants of an evening training course at Eötvös Loránd University, and their proficiency was upper-intermediate. The one-way analysis of variance of the C-test scores showed that the level of proficiency of the three groups was significantly different ($F = 130.45$; $p < 0.001$).

Both the speech produced while performing the task and the text of the retrospective interview were recorded and transcribed by trained research assistants. Transcriptions were checked by the researcher.

4.2. Analysis

The self-repairs were identified and classified in the texts with the help of the retrospective interviews. Four major types of self-repairs were established with sub-categories within each group. **Different information (D-) repairs** involve the decoding of different information than the speaker is currently formulating. **Appropriacy (A-) repairs** are used when the speaker modifies the originally intended information. **Error repairs** involve corrections of accidental lapses in accessing words (**lexical error repairs**), reprocessing the grammatical structure of the utterance (**grammatical error repairs**) and revised phonological encoding mechanisms (**phonological error repairs**). **Rephrasing-repairs** are used by L2 speakers when they are uncertain about the correctness of the original utterance, and in this case they encode their original message in a different form. (For subtypes of D- and A-repairs as well as for examples see the appendix).

The analysis of well-formedness employed in the study differed slightly from previous research in this field. If one considers Levelt's (1983) rule presented above, it does not specify how one should proceed in the case of within-word interruptions, as it only says that there should be a

string of zero or more words to complete the reparandum. In the case of within-word interruptions, however, it is often a fragment of a word (and frequently not even a well-formed syllable) that is needed to finish the original utterance. Due to the fact that in these cases it can only be speculated how speakers intended to utter their message, no accurate judgements can be made concerning the grammaticality and well-formedness of repairs with within-word interruptions. Therefore, these types of repairs ($N = 33$, 11.4% of all the repairs in the corpus in the L2 corpus and $N = 8$, 16.3% in the L1 corpus) were excluded from the analysis of well-formedness in this study. The L1 corpus consisted of 59 and the L2 corpus of 289 self-repairs altogether. It has to be noted here that the number of L1 self-repairs was not high enough to make overall generalizations concerning the structure of self-corrections in Hungarian, but it was sufficient as background to the L2 data.

Editing terms preceding the reparatum were all identified together with the point of interruption of the reparandum in the transcripts. For the analysis of the effect of proficiency on the structure of self-repairs one-way analysis of variance and Chi-square analysis and for the investigation of the relationship of the type and structure of repair Chi-square analysis was used. The level of significance was set for $p < 0.05$ in the study.

5. Results and discussion

5.1. The well-formedness of repairs

As can be seen in Table 1, 87.3% of the repairs in the L2 data-base were well-formed. The highest frequency of ill-formed repairs could be found within the subtype of grammatical error-repairs (see Example 3 above) and message-replacement repairs (i.e., repairs when the speaker abandons the original speech plan and encodes a completely new message). Due to the high number of cells with values lower than 5, no meaningful Chi-square statistics could be computed with subtypes of repairs. The analysis of the relationship of main types of repairs, however, showed that the type of correction significantly influences whether the correction is well-formed ($\chi = 13.78$, $p = 0.003$). Error-repairs were ill-formed more frequently than expected, whereas appropriacy repairs were more frequently well-formed than expected.

Table 1

The distribution of well-formed and ill-formed repairs in L2

Type of repairs	WELL-FORMED Count	Expected value	Row per-cent. (%)	Residual	ILL-FORMED Count	Expected value	Row per-cent. (%)	Residual
EL-repairs	33	28.8	100.0	4.2	0	4.2	0	−4.2
EG-repairs	26	38.4	59.1	−12.4	18	5.6	40.9	12.4
EF-repairs	8	7.9	88.9	0.1	1	1.1	11.1	−0.1
AL-repairs	39	34.9	97.5	4.1	1	5.1	2.5	−4.1
AA-repairs	3	2.6	100	0.4	0	0.4	0	−0.4
ALC-repairs	2	1.7	100	0.3	0	0.3	0	−0.3
AP-repairs	10	8.7	100	1.3	0	1.3	0	−1.3
AG-repairs	7	6.1	100	0.9	0	0.9	0	−0.9
DM-repairs	34	34.9	85.0	−0.9	6	5.1	15.0	0.9
DI-repairs	16	14.8	94.1	1.2	1	2.2	5.9	−1.2
DO-repairs	3	2.6	100	0.4	0	0.4	0	−0.4
R-repairs	38	37.5	88.4	0.5	5	5.5	11.6	−0.5
Rest	4	3.5	100.0	0.5	0	0.5	0	−0.5
Total	**219**		**87.3**		**32**		**12.7**	

EL-repairs = lexical error-repairs, EG-repairs = grammatical error-repairs; EF-repairs = phonological error repairs; AL-repairs = appropriate level of information-repairs; AA-repairs = ambiguous reference-repairs; ALC-repairs = coherent terminology repairs; AP-repairs = pragmatic appropriacy repairs; AG-repairs = repairs for good language; D-repairs = different information repairs; DM-repairs = message replacement-repairs; DI-repairs = inappropriate information repairs; DO-repairs = ordering error-repairs; R-repairs = rephrasing-repairs

The Chi-square analysis of the relationship of proficiency grouping and well-formedness showed a significant effect of foreign language competence ($\chi = 6.69$, $p = 0.03$). In the intermediate group ill-formed repairs occurred with a higher frequency than expected whereas in the speech of advanced learners there were more well-formed repairs than expected (see Table 2, overleaf). As for the effect of language, the results showed that the percentage of well- and ill-formed repairs did not differ significantly in Hungarian and in English as L2 ($\chi = 1.15$, $p = 0.28$).

In the Hungarian data-base 73.2% of the repairs were well-formed and 26.8% ill-formed. The Chi-square analysis showed a significant relationship between the main type of repair and well-formedness ($\chi = 11.78$, $p = 0.002$). However, 33.3% of the cells did not have an expected frequency above 5, which shows that the corpus is too small for drawing a firm conclusion on the basis of the findings. The results indicate

Table 2

The relationship of well-formedness and the level of proficiency in L2

Level of proficiency (count, expected, row %, residual)	Ill-formed	Well-formed
Intermediate	16	64
	10.2	69.8
	20.0%	80.0%
	5.8	−5.8
Upper-intermediate	10	72
	10.5	71.5
	12.2%	87.8%
	−0.5	0.5
Advanced	6	83
	11.3	77.7
	6.7%	93.3%
	−5.3	5.3
Total	**32**	**219**
	12.7%	**87.3%**

$\chi = 6.69$, $p = 0.03$

that different-information repairs in Hungarian were more frequently ill-formed than expected.

The findings suggest that in most of the cases, the self-repair behaviour of Hungarian learners and native speakers is also governed by the well-formedness rule. Thus well-formedness seems to be a phenomenon independent of language background. This lends additional support to Kormos's hypothesis that both L1 speakers and L2 learners "are able to store the original syntactic structure of their message in working memory" (1999, 329) and adjust the reparatum to it. The hypothesis that both native and non-native speakers implement the correction by reproducing the syntactic environment of the reparandum, that is, "by grammatically encoding the relevant part of the message anew" (*idem.*) also seems to be supported. The findings of the study lend strong support to the assumption that speech production usually does not start from the intermediary levels where the error was made (Kormos 1999), thus the results indirectly support modular models of speech production (e.g., Levelt 1989; Levelt et al. 1999) and Levelt's (1989) model of monitoring. In this the-

ory three monitor loops are assumed to be responsible for inspecting the outcome of the production processes. The first loop involves the comparison of the preverbal plan with the original intentions of the speaker. The second loop concerns the monitoring of the phonetic plan (i.e., 'internal speech') before articulation, which is also called 'covert monitoring' (see also Postma–Kolk 1992, 1993; Wheeldon–Levelt 1995). Finally, the generated utterance is also checked after articulation, which constitutes the final, external loop of monitoring, involving the acoustic-phonetic processor. Upon perceiving an error or inappropriacy in the output in any of these three loops of control, the monitor issues an alarm signal, which, in turn, triggers the production mechanism for a second time starting from the phase of conceptualization.

The findings of this study also suggest that well-formedness seems to be a universal phenomenon in speech production. This does not necessarily mean that this study lends support to theories claiming that L2 speakers have access to Universal Grammar (Chomsky 1965). It is rather the case that language users in general aim to produce well-formed sentences, otherwise their interlocutors have difficulty understanding them. Therefore, even if the utterance contains a slip, error or inappropriacy, speakers will strive to adjust the syntactic structure of the newly formulated message to the preceding part of the utterance. Since certain maxims of conversation are supposed to be universal such as the maxim of clarity (Grice 1975), speakers—regardless of L1 and independent of the fact whether they speak their L1 or another language—will aim to make their message clear and unambiguous.

In comparison with previous research that investigated the syntactic structure of self-repairs, it can be stated that the rate of well-formed self-corrections in this project is between the percentages in Levelt's (1983) (L1: 98%) and van Hest's (1996) (L1: 70% and L2: 80%) studies. Van Hest explained the lower rate of well-formed repairs found in her corpus with reference to the fact that in her study the task to be performed by the students was more complex than in that of Levelt (1983). She argued that, for this reason, participants of her project were required to use syntactically varied sentences, which resulted in a relatively low percentage of well-formed repairs. Accordingly, van Hest assumed that the higher rate of well-formed repairs in L2 was due to the fact that utterances of speakers are more complex in L1 than in L2, therefore, self-corrections in L2 are syntactically easier to implement than those in L1. It has to be noted that van Hest only used descriptive statistics in

her study, therefore it is not known whether the difference in the rate of well-formed repairs between L1 and L2 was statistically significant. In the present study well-formedness was not significantly influenced by whether the participant spoke L1 and L2. In addition, van Hest's line of argumentation would also mean that with increasing proficiency L2 speakers make a decreasing number of well-formed self-repairs, which seems to contradict the results of this study where an opposite effect was found.

The finding that 20% of the self-repairs produced by intermediate learners was ill-formed as opposed to the 6.7% of ill-formed repairs in advanced learners' speech can be explained with reference to the psycholinguistic mechanism of self-correction and theories of attention. From psychological research it is well known that attention is limited due to the constraints of working memory (Gathercole–Baddeley 1994). It is also an accepted fact that the less automatic a mechanism is, the more attention is required for it (for a review see Schmidt 1992). Thus it seems that intermediate learners need to pay so much attention to other aspects of encoding their message anew that they are often unable to keep the syntactic structure of their original message in their working memory. In turn, this frequently results in ill-formed self-corrections.

As regards the varying percentage of well-formed and ill-formed repairs in the case of different subtypes of self-repairs in L2, it is understandable that grammatical error-repairs do not always follow the well-formedness rule. When implementing changes in the syntactic structure of the message (e.g., in the word order), the coordination of the same phrasal category is often impossible. When producing message-replacement repairs, L2 speakers need to conceptualise a completely new message instead of the one to be replaced. The structure of this new utterance might not follow the previous one due to the lack of attentional resources, as learners might be so overloaded with the cognitive planning of the repair that they have no attention to spare to adjust the syntactic structure of the reparatum to that of the reparandum. Results concerning the time necessary for re-planning message-replacement repairs also support the increased cognitive load in this case (Kormos 2000b). In the case of L1 self-corrections, the corpus is too small to observe any meaningful relationship between the type of repair and well-formedness.

5.2. The use of editing terms

With respect to the editing terms used upon making the repair, the research project investigated two important issues: the relationship of the types of self-correction and the editing terms used by the participants as well as the effect of language proficiency on the frequency and type of editing terms.

As can be seen in Table 3 (overleaf), the results of the research project replicated previous findings concerning the default repair structure in L2, since 55% of the self-corrections contained no editing term. Similarly to earlier studies in this field, it was found that in the majority of the cases when editing expressions were used, the repair was merely signalled by the presence of filled non-lexicalized pauses, such as *er* or *uhm*. Among lexicalized editing terms *or* was most frequently applied, and interestingly, Hungarian expressions were very rarely used for this purpose. Certain editing terms tended to cooccur with certain types of repairs. *Sorry* was employed exclusively with error-repairs, whereas *I mean* always indicated appropriacy or different information repairs.

The relationship of the main types of self-corrections and editing terms was also studied by means of Chi-square analysis. The results of this investigation indicated that certain types of repairs cooccurred with specific editing terms ($\chi^2 = 19.47$, $p = 0.0001$). Error-repairs contained unfilled pauses with higher frequency than lexicalized editing terms, while different information-repairs tended to be signalled by lexicalized expressions more often than by unfilled pauses.

Editing terms were used differently in Hungarian than in English. Interestingly the default structure of repairs in Hungarian was not the lack of editing term, but the use of the term *tehát* 'so'. Due to the small number of L1 self-repairs, no meaningful Chi-square statistics could be computed concerning the relationship of types of repairs and editing terms in Hungarian.

The influence of the level of proficiency on the frequency of the use of editing terms was also analysed by means of one-way analysis of variance. The results of the study indicate limited effect of proficiency in this respect. The frequency of the editing terms was not significantly affected by the participants' level of competence in English. It could be observed, however, that certain editing terms such as *I mean* or *well* were merely applied by highly proficient speakers, while Hungarian editing terms occurred only in the speech of participants with low level of competence.

Table 3

The distribution of editing terms across different types of self-repairs in L2

Type of repairs (count, row %)	Unfilled pause	er, uhm	so	or	sorry	I mean	and	well	other	ja (ugh)	vagy (or)
EL-repairs	22	10	0	2	4	0	0	0	1	0	2
	53.7	24.4	0	4.9	9.8	0	0	0	2.4	0	4.9
EG-repairs	33	15	0	0	0	0	1	0	0	0	0
	67.3	30.6	0	0	0	0	2.0	0	0	0	0
EF-repairs	19	3	0	0	0	0	0	0	0	0	0
	11.9	3.8	0	0	0	0	0	0	0	0	0
AL-repairs	21	8	4	7	0	2	0	0	0	0	0
	50.0	19.0	9.5	16.7	0	4.8	0	0	0	0	0
AA-repairs	2	1	0	0	0	0	0	0	0	0	0
	66.7	33.3	0	0	0	0	0	0	0	0	0
ALC-repairs	1	1	0	0	0	0	0	0	0	0	0
	50.0	50.0	0	0	0	0	0	0	0	0	0
AP-repairs	8	2	0	0	0	0	1	0	0	0	0
	72.7	18.2	0	0	0	0	9.1	0	0	0	0
AG-repairs	6	1	0	0	0	1	0	0	0	0	0
	75.0	12.5	0	0	0	12.5	0	0	0	0	0
DM-repairs	18	18	2	1	0	2	0	1	1	0	0
	41.9	41.9	4.7	2.3	0	4.7	0	2.3	2.3	0	0
DI-repairs	6	1	0	5	0	0	0	0	4	1	0
	35.3	5.9	0	29.4	0	0	0	0	23.5	5.9	0
DO-repairs	1	1	0	0	0	0	0	1	0	0	0
	33.3	33.3	0	0	0	0	0	33.3	0	0	0
R-repairs	22	16	1	2	0	0	1	1	1	0	0
	50.0	36.4	2.3	4.5	0	0	2.3	2.3	2.3	0	0
Total	**159**	**80**	**7**	**18**	**4**	**5**	**3**	**3**	**7**	**1**	**2**
	55.0	**27.7**	**2.4**	**6.2**	**1.4**	**1.7**	**1.0**	**1.0**	**2.4**	**0.3**	**0.7**

E-repairs = error repairs; EL-repairs = lexical error-repairs, EG-repairs = grammatical error-repairs; EF-repairs = phonological error-repairs; A-repairs = appropriacy-repairs; AL-repairs = appropriate level of information-repairs; AA-repairs = ambiguous reference-repairs; ALC-repairs = coherent terminology repairs; AP-repairs = pragmatic appropriacy repairs; AG-repairs = repairs for good language; D-repairs = different information repairs; DM-repairs = message replacement-repairs; DI-repairs = inappropriate information repairs; DO-repairs = ordering error-repairs; R-repairs = rephrasing-repairs

The findings of the present study lend strong support to previous assumptions concerning Dutch and English repairs, namely, that in most of the cases speakers do not use any signals except for an unfilled pause for indicating that a repair will be made. One of the reasons for this might be

that making repairs hinders fluent expression, and the use of long editing terms would make the flow of speech even more disrupted. Another possible explanation is that in most of the cases when a correction is implemented, it is obvious that the speaker intends the listener to ignore the previously uttered message (e.g., in the case of non-existing words and grammatical errors); thus, there is no need to signal it. This seems to be especially true in the case of L2 error-repairs, which are most frequently accompanied by unfilled pauses in the present corpus.

With respect to different-information repairs in L2, however, listeners frequently need to be warned that the message to be conveyed next might not logically follow the previously uttered one. For example, in cases when part of a message is totally replaced by a new one (message-replacement repairs), or when the previously provided information has been false (inappropriate information repair), the listeners' activated schemata of what to expect as a continuation can be disturbed by the repair. Thus, speakers might perceive that a warning signal is more essential upon producing these types of corrections than in cases when it is the reparandum (e.g., the slip of the tongue) and not the reparatum which does not fit the expectations of the listener. The use of editing terms in these cases is very similar to that of **verbal strategy markers**, which can stand "before or after a (communication) strategy to signal that the word or structure does not carry the intended meaning perfectly in the L2 code" (Dörnyei–Scott 1997, 191). The similarity of these two markers lies in the fact that they both aim to elicit cooperation from the interlocutor and to achieve that the listener and the speaker share the same meaning.

The results concerning the cooccurence of different types of editing terms and repairs in L2 indicate that the participants of the study use these expressions in a similar way as native speakers of English do. They are aware of the fact that *I mean* is applied for further specifying the informational content of the message, and apologising (*sorry*) is only necessary after having made an error (DuBois 1974; cf. Levelt 1989). Moreover, these results also indicate that the use of editing terms is not random, but systematic both in L1 and L2 speech. Most editing terms including filled pauses tend to cooccur with specific types of repairs, that is, they signal a given reason for communication breakdown. This supports Clark's and Fox Tree's (2002) proposal concerning the process of self-corrections, in which they claim that speakers consciously and systematically select the signal for making a repair.

The infrequent use of L1 editing terms by the participants of this project seems to contradict Poulisse and Bongaerts' (1994) and van Hest's (1996) findings. In both of these studies beginning and intermediate learners transferred editing expressions from L1 to L2 considerably often, whereas in the present corpus L1 editing terms occurred only in the speech of intermediate learners and with very low frequency. One of the possible explanations for this difference might be that due to the fact that the L1 of the participants of both Poulisse and Bongaert's (1994) and van Hest's (1996) study was Dutch. Dutch learners of English might be more willing to transfer the editing expressions from their L1 to L2 than Hungarian speakers, whose mother tongue has few characteristic features in common with English (for the effect of the proximity of languages on transfer see Kellerman 1979).

The effect of language proficiency on the use of editing terms seems to be limited, the reason for which can be that it is rather the nature of the repair that determines what type of editing expression will be used than the level of L2 competence. Nevertheless, it can be observed that only advanced learners apply *well* and *I mean* for signalling the correction, which is probably the consequence of the lack of instruction of discourse markers in most of the language courses in Hungary.

It is interesting to observe that despite the fact that there were very few error repairs in the Hungarian corpus, the speakers used a lexical editing term in 69.4% of the cases and that Hungarian speakers used the editing term *tehát* 'so' the most frequently. Due to the small number of participants and self-corrections, one can only speculate that perhaps Hungarian speakers prefer to signal corrections with the term *tehát* 'so' rather than simply to use a non-filled pause.

5.3. The point of interruption

The point of interruption was also investigated both in L2 and in L1 of the participants. As can be seen in Table 4, interruptions within the trouble-word occurred infrequently in the L2 corpus of self-repairs of the present study. The highest frequency of within-word cut-off points could be found in the case of phonological error-repairs, and words containing lexical or grammatical errors were also frequently interrupted. The results suggest that the flow of speech is mainly halted after the reparandum has been uttered. Late interruptions were rare in the corpus.

Table 4

The placement of cut-off points in the case of various sub-types of self-repairs in L2

Type of repairs (count, row %)	Within the trouble word	Immediately following the trouble word	Within the word following the trouble word	1 word after the trouble word	2 words after the trouble word	Later than 2 words after the trouble word
EL-repairs	7	23	2	5	2	1
	17.5	57.5	5.0	12.5	5.0	2.5
EG-repairs	6	32	1	3	6	2
	12.0	64.0	2.0	6.0	12.0	4.0
EF-repairs	13	8	0	1	0	0
	59.1	36.4	0.0	4.5	0.0	0.0
ALC-repairs	2	32	1	3	2	2
	4.8	76.2	2.4	7.1	4.8	4.8
AA-repairs	0	3	0	0	0	0
	0.0	100.0	0.0	0.0	0.0	0.0
AC-repairs	0	1	0	1	0	0
	0.0	50.0	0.0	50.0	0.0	0.0
AP-repairs	1	8	1	1	0	0
	9.1	72.7	9.1	9.1	0.0	0.0
AG-repairs	1	5	0	1	1	0
	12.5	62.5	0.0	12.5	12.5	0.0
DM-repairs	2	38	1	2	0	0
	4.7	88.4	2.3	3.0	0.0	0.0
DI-repairs	0	13	0	1	2	1
	0.0	76.5	0.0	5.9	11.8	5.9
DO-repairs	0	3	0	0	0	0
	0.0	100.0	0.0	0.0	0.0	0.0
R-repairs	1	34	1	2	3	3
	2.3	77.3	2.3	4.5	6.8	6.8
Total	**33**	**200**	**7**	**20**	**16**	**9**
	11.6	**70.2**	**2.5**	**7.0**	**5.6**	**3.1**

EL-repairs = lexical error-repairs, EG-repairs = grammatical error-repairs; EF-repairs = phonological error-repairs; AL-repairs = appropriate level of information-repairs; AA-repairs = ambiguous reference-repairs; ALC-repairs = coherent terminology repairs; AP-repairs = pragmatic appropriacy repairs; AG-repairs = repairs for good language; D-repairs = different information repairs; DM-repairs = message replacement-repairs; DI-repairs = inappropriate information repairs; DO-repairs = ordering error-repairs; R-repairs = rephrasing-repairs

As the number of cells with values lower than 5 was high when all the types and possible cut-off points were computed, meaningful Chi-square statistical analyses concerning the interruption pattern of the different

main groups of self-repairs in L2 could only be performed when the cut-off points were divided into three categories: (1) interruptions within the trouble word, (2) interruptions immediately following the trouble word, and (3) interruptions later than following the trouble word. The main groups of self-repairs varied as regards the placement of the cut-off points to a significant extent ($\chi^2 = 29.97$, $p = 0.0001$). The results indicate that erroneous words were interrupted with a higher frequency than expected, whereas inappropriacies, words containing false information, and lexical items in the correctness of which speakers were uncertain tended to be completed.

As error-repairs were found to behave in a different manner concerning the placement of cut-off points, the interruption pattern of each subtype of this group of corrections was also analysed by means of Chi-square statistics. The results indicate that it is mainly the class of phonological error-repairs that contributed to the above observed differences, as words containing phonological errors were more frequently interrupted than the expected value ($\chi^2 = 21.68$, $p = 0.0001$). As regards the other two types of error-repairs, no considerable deviations from the usual distribution of cut-off points could be observed. The level of proficiency was not found to affect the placement of cut-off points ($\chi^2 = 1.07$; $p = 0.89$).

In the Hungarian corpus 71.4% of the repairs were interrupted immediately after the trouble word, 16.3% of the repairs were within-word interruptions, in 10.2% of the repairs the interruption took place one word after the trouble word and in 2% of the cases two words after the trouble word. Due to the small number of repairs, no meaningful Chi-square statistics could be computed concerning the relationship of the type of repair and the point of interruption.

The placement of cut-off points in relation to word boundaries is an unreliable indicator of detection processes, therefore, this aspect will not be discussed here. Thus, the conclusions one can draw from the results concerning the interruption pattern of different types of repairs are limited to the structural and pragmatic aspects of this issue.

Similarly to previous studies in this field (e.g., Brédart 1991; van Hest 1996), the findings of the present project suggest that Hungarian learners of English also interrupt erroneous words, while inappropriate lexical items tend to be completed. This can be explained with reference to the fact that erroneous words need to be ignored by the listener when decoding the interlocutor's message, consequently, they tend to be interrupted. Inadequate information, however, only needs further speci-

fication, and as words containing inappropriacies often carry meaning for the listener (Brédart 1991), they will be completed. No such tendency could be observed in the case of Hungarian native speakers, which might be due to the small number of repairs in the corpus. The findings concerning L2 self-repairs indicate that not only the choice of editing terms and lexical fillers is conscious and systematic but also decisions concerning the interruption of the flow of speech when a correction is needed. This shows that the process of monitoring and self-repair is a conscious process, which is governed by universal pragmatic constraints that require communication to be clear and unambiguous.

6. Conclusion

The investigation of the structure of self-repairs in the speech of Hungarian learners of English yielded a number of results which confirm earlier studies on the monitoring behaviour of L2 speakers from a Dutch native speaking background. From the analysis of the well-formedness of self-repairs, it was concluded that both self-corrections in Hungarian and in the speech of Hungarian learners adhere to the rules of syntactic coordination. This implies that in most cases L2 speakers also store the syntactic structure of the reparandum in their working memory and adjust the reparatum to it. In addition, these findings also support earlier assumptions that the majority of self-repairs do not involve the mere substitution of one lexical entry with another one, but the encoding of the relevant part of the speech plan completely anew. Therefore, the study lends additional support for modular models of speech production (e.g., Levelt 1989; Levelt et al. 1999). A new finding of the research was that the level of proficiency has a significant effect on the rate of well-formed repairs because intermediate learners produced fewer well-formed repairs than advanced students. This was explained with reference to the limited attentional resources of low level learners.

The findings concerning the use of editing terms in L2 also confirmed the assumptions of previous studies, namely, that the default repair structure contains no editing term, and if editing terms are used, they are most frequently unfilled pauses. L2 speakers only signal the repair with a lexical editing term if they assume that the reparatum might not fit the listener's schemata activated by the reparandum, for example in the case of the abandonment of the original message and when the original information provided was inappropriate. In this case editing terms

can also be interpreted as verbal strategy markers, which indicate that the speaker needs the interlocutor's attentive cooperation (Dörnyei–Scott 1997). When L2 speakers presume that their listeners will easily perceive that the previously uttered message is erroneous (e.g., because it contains a non-existing word or a grammatical error), and it needs to be ignored, they do not signal that a repair will be made, as the use of an editing term would make the flow of speech even more disruptive. The fact that the level of L2 competence was not found to affect the frequency of the use of various editing terms also supports the assumption that the major factor that determines the use of editing terms is the type of repair, that is, the nature of error. The interruption pattern of errors and inappropriacies upon making the repair in L2 was found to show analogous tendencies. Erroneous words that speakers intended to be ignored tended to be interrupted more frequently than parts of utterances conveying inadequate information as the latter types of reparandum could carry meaning for the listener and needed only further specification.

The results of the study indicate that it is not only the process of monitoring that is conscious and controlled in nature, but speakers also make conscious and systematic decisions concerning the interruption of the flow of speech and the selection of editing terms. These decisions are primarily governed by the Gricean (1975) maxim of clarity, which requires that utterances in a conversation should be clear and unambiguous. This constraint involves considering the communicative situation and context and the interlocutor's perspective, which shows a strong relationship between pragmatic and discourse knowledge and psycholinguistic processing.

Appendix

The classification of self-repairs (quoted from Kormos 2000a, 380–3)

Name of repair	Definition	Example
Different information (D-) repair	The speaker decides to encode different information from the one he/she is currently formulating (Levelt 1983)	
Inappropriate information (DI-) repair	The speaker repairs the message because its information content is faulty (Levelt 1983)	The room is er uhm eer thirty thirty thousand er too much er ten thousand er forint er forints per day

Ordering error (DO-) repair	The decides to encode parts of the intended message in different order (Levelt 1983)	Well, we it's it's about a thousand Forints Retrospection: First I wanted to answer the second question, but then I realized that I should answer the first question first.
Message abandonment (DM-) repair	The speaker abandons the originally intended message and replaces it with a different one	we have some er er v... maybe you have vegetarians in your group Retrospection: Here the idea of vegetarians suddenly popped up, and I abandoned what I was going to say because I would not have been able to list any more types of food anyway.
Appropriacy (A-) repair	The speaker decides to encode the original information in a modified way (Levelt 1983)	
Appropriate level of information (AL-) repair	The speaker decides to further specify the original message (Levelt 1983).	There are very wide choice of er main courses er er steak er er several kind of steak Retrospection: I wanted to say it more precisely that we do not only have one kind of steak but several kinds of steak.
Ambiguous reference (AA-) repair	The speaker repairs the referring expression because of ambiguity (Levelt 1983).	And you have to pay extra for the drinks. Then you have to negotiate that and talk about the drinks with the barman. Retrospection: I corrected what I said because it was not clear whether you have to talk about the drinks or the price with the barman.
Coherent terminology (AC-) repair	The speaker repairs incoherent terminology (Levelt 1983).	in this case er if it is so urgent and important for you, we would like er you to:: to write us **an order**—er in er 24 hours that you make sure that you will er come and book this eel room. ... R: I see, all right and then I can only pay the deposit next week when I er find out how many people come and when I have talked to all of the people. S: Er but **this letter** is er— **the order** —er your request is er anyway—needed and we:::

		Retrospection: I remembered that I had used the word 'order' earlier, and I wanted to stick to the same terms, so I replaced 'letter' with 'order'.
Pragmatic appropriacy (AP-) repair	The speaker repairs part of the message which is pragmatically inappropriate in the given situation (based on Brédart 1991)	Can I what can I do for you? Retrospection: First I wanted to say 'can I help you', but I thought this is said in shops only, and I decided to say 'what can I do for you' because it was more appropriate in this situation.
Repair for good language (AG-)	The speaker repairs part of the message which he/she judges to be not sophisticated enough concerning the manner of expression	If you want the room, I mean if you decide on it Retrospection: I was not satisfied with this sentence with the word 'want' in it, I did not like it stylistically.

References

Bot, Kees de 1992. A bilingual production model: Levelt's 'speaking' model adapted. In: Applied Linguistics 13 : 1–24.

Brédart, Simon 1991. Word interruption in self-repairing. In: Journal of Psycholinguistic Research 20 : 123–37.

Chomsky, Noam 1965. Aspects of the theory of syntax. MIT Press, Cambridge MA.

Clark, Herbert H. – Jean E. Fox Tree 2002. Using *uh* and *um* in spontaneous speaking. In: Cognition 84 : 73–111.

Dell, Gary S. 1986. A spreading activation theory of retrieval in sentence production. In: Psychological Review 93 : 283–321.

Dell, Gary S. – Peter O'Seaghda 1991. Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). In: Psychological Review 98 : 604–14.

Dörnyei, Zoltán – Lucy Katona 1992. Validation of the C-test amongst Hungarian EFL learners. In: Language Testing 9 : 187–206.

Dörnyei, Zoltán – Mary Lee Scott 1997. Communication strategies in a second language: Definitions and taxonomies. In: Language Learning 47 : 173–210.

DuBois, John 1974. Syntax in mid-sentence. In: Charles Fillmore – George Lakoff – Robert Lakoff (eds): Berkeley studies in syntax and semantics III.1–III.25. University of California, Institute of Human Learning and Department of Linguistics, Berkeley CA.

Ericsson, K. Anders – Herbert A. Simon 1980. Verbal reports as data. In: Psychological Review 87 : 215–251.

Ericsson, K. Anders – Herbert A. Simon 1993. Protocol Analysis: Verbal reports as data. MIT Press, Cambridge MA.

Fromkin, Victoria (ed.) 1980. Errors in linguistic performance: Slips of the tongue, ear, pen and hand. Academic Press, New York.

Fry, Dennis 1969. The linguistic evidence of speech errors. In: Brno Studies of English 8 : 69–74.

Garrett, Merrill F. 1976. Syntactic processes in sentence production. In: Roger J. Wales – Edward Walker (eds): New approaches to language mechanisms, 274–93. North Holland, Amsterdam.

Gathercole, Susan E. – Alan Baddeley 1994. Working memory and language. Hillsdale NJ, Lawrence Erlbaum.

Gósy, Mária (ed.) 2004. Beszédkutatás 2004. „Nyelvbotlás"-korpusz tanulmányok [Speech research 2004. Studies of the "slip of the tongue" corpus]. MTA Nyelvtudományi Intézet, Budapest.

Grice, H. Paul 1975. Logic and conversation. In: Peter Cole – Jerry L. Morgan (eds): Syntax and semantics, vol. 3: Speech acts, 41–58. Academic Press, New York.

Hebb, Donald O. 1949. The organization of behavior. New York, Wiley.

Hest, Erna van 1996. Self-repair in L1 and L2 production. Tilburg University Press, Tilburg.

Huszár, Ágnes 2005. A gondolattól a szóig. A beszéd folyamata a nyelvbotlások tükrében [From thoughts to words. The process of speech production in the light of slips of the tongue]. Tinta Könyvkiadó, Budapest.

Kellerman, Eric 1979. Transfer or non-transfer: Where are we now? In: Studies in Second Language Acquisition 2 : 37–57.

Kormos, Judit 1999. Monitoring and self-repair in L2. In: Language Learning 49 : 303–42.

Kormos, Judit 2000a. The role of attention in monitoring second language speech production. In: Language Learning 50 : 343–84.

Kormos, Judit 2000b. The timing of self-repairs in second language speech production. In: Studies in Second Language Acquisition 22 : 145–69.

Laver, John 1980. Monitoring systems in the neurolinguistic control of speech production. In: Fromkin (1980, 287–305).

Levelt, Willem. J. M. 1983. Monitoring and self-repair in speech. In: Cognition 33 : 41–103.

Levelt, Willem. J. M. 1989. Speaking: From intention to articulation. MIT Press, Cambridge MA.

Levelt, Willem J. M. – Ardi Roelofs – Antje S. Meyer 1999. A theory of lexical access in speech production. In: Behavioural and Brain Sciences 22 : 1–75.

Nooteboom, Sieb G. 1980. Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech. In: Fromkin (1980, 87–95).

Postma, Albert – Herman Kolk 1992. The effects of noise masking and required accuracy on speech errors disfluencies and self-repairs. In: Journal of Speech and Hearing Research 35 : 537–544.

Postma, Albert – Herman Kolk 1993. The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. In: Journal of Speech and Hearing Research 36 : 472–487.

Poulisse, Nanda 1999. Slips of the tongue. Speech errors in first and second language production. John Benjamins, Amsterdam & Philadelphia.

Poulisse, Nanda – Theo Bongaerts 1994. First language use in second language production. In: Applied Linguistics 15 : 36–57.

Schmidt, Richard 1992. Psychological mechanisms underlying second language fluency. In: Studies in Second Language Acquisition 14 : 357–85.

Smedt, Konrad de – Gerard Kempen 1987. Incremental sentence production, self-correction, and coordination. In: Gerard Kempen (ed.): Natural language generation: Recent advances in artificial intelligence, psychology and linguistics, 365–76. Kluwer, Dordrecht.

Stemberger, Joseph P. 1985. An interactive activation model of language production. In: Andrew W. Ellis (ed.): Progress in the psychology of language. Vol. 1, Hillsdale NJ. Lawrence Erlbaum, 143-86.

Wheeldon, Linda R. – Willem J. M. Levelt 1995. Monitoring the time course of phonological encoding. In: Journal of Memory and Language 34 : 311–34.

Acta Linguistica Hungarica, Vol. 53 (1), pp. 77–94 (2006)
DOI: 10.1556/ALing.53.2006.1.4

BOOK REVIEWS

Katalin É. Kiss: The syntax of Hungarian. Cambridge University Press, Cambridge, 2002, 278 pp.

The book represents a systematic study of Hungarian sentence structure. Throughout the book É. Kiss uses a large amount of results from several linguists. Yet, it appears to repeat the basic proposal made already in the seventies by É. Kiss, namely, that the Hungarian sentence consists of a topic and a predicate. The predicate is a head initial verb phrase, and it can be preceded by focused elements and quantifiers.

The results of other researchers on Hungarian syntax are adopted or rather adapted to this basic premise. This results in several cases in analyses where the flat VP structure is retained and the proposals of other authors are slightly altered in order to fit É. Kiss's general framework.

The book under review has the following structure. Chapter 1 is an introductory chapter that presents the necessary background assumptions and some information about Hungarian in general, its distribution, genealogy and an overview of syntactic and morphosyntactic features. A short summary of the next chapters is also included in the introduction which helps to guide the reader. Chapters 2 to 6 are devoted to operator positions that are projected above the flat VP projection. Chapter 2 discusses the topic phrase, the function of topics and the operation of topicalization. Chapter 3 deals with the predicate phrase itself and examines the relation of arguments within the VP and the properties of the verbal modifier (VM). Chapter 4 describes the semantic and syntactic properties of the focused element. Chapter 5 investigates quantifier phrases, their position, scope and interpretation. Chapter 6 deals with negative phrases, negation and negative concord. Chapter 7 deals with the structure of the noun phrase. It is claimed that the inner structure of the various types of complements parallels the inner structure of the extended verb phrase. The noun phrase also consists of a lexical kernel and it is further extended by operator and morphosyntactic projections like AgrP. Chapter 8 discusses the structure of the postpositional phrase

and Chapter 9 is devoted to non-finite verbal projections. Non-finite verbal projections are shown to be also extended by morphosyntactic and operator projections. The parallels between finite and non-finite verbal projections are emphasized throughout the chapter. The last chapter is dedicated to the internal structure of subordinate clauses which is claimed to be basically identical to the structure of matrix clauses.

In each chapter the reader is led through detailed argumentations, a wealth of data and presentations of several previous analyses with their merits and shortcomings discussed. In those cases where É. Kiss adopts analyses from other researchers the source and the motives are clearly stated although sometimes the presentation of the adopted analysis is rather concise.

Chapter 2 deals with the topic phrase and the properties of the topic. First, É. Kiss gives a definition of the topic function: "The topic foregrounds an individual from among those present in the universe of discourse as the subject of the subsequent predication." The formal features of a topic constituent are examined and it is concluded that the topic constituent must be referential and specific. É. Kiss then examines how the topic-predicate boundary can be located. The most obvious clue is stress, since in Hungarian the obligatory stress falls on the first major constituent of the predicate phrase. Therefore the topic cannot bear major stress. Sentence adverbials can precede or follow the topic, but can never enter the predicate phrase. After presenting the empirical data about Hungarian sentences with a topic phrase, É. Kiss proposes to analyze the topic constituent as an argument of the verb that has been preposed from the VP and binds an argument position in it. The topic constituent moves to a functional projection called topic phrase (TopP). This TopP can be iterated. Topicless sentences are also examined in this chapter and it is claimed that a sentence can be topicless if it involves a logical propositional operator. Such sentences can be both stative and eventive.

A rather extensive chapter, Chapter 3, is dedicated to the structure of the minimal predicate. It consists of a VP, extended by morphosyntactic projections such as modality, tense, mood and agreement and further extended into an aspectual phrase.

A very long subsection of Chapter 3 deals with the morphosyntactic projections that extend the core VP. Hungarian is a agglutinative language, its tense, mood and agreement morphemes appearing as suffixes on the verb. In this section É. Kiss basically adopts the analysis of Bartos (1999). Bartos analyzes these morphosyntactic suffixes as independent syntactic constituents occupying head positions of functional projections. Bartos extends the verb phrase with five further functional projections.

The final issue dealt with in Chapter 3 is the category and structural position of a particle-like adverbial element traditionally called the verbal prefix. The analysis of the verbal prefix presents several difficult questions. First, the prefix + verb combination shows characteristics of a lexical unit that should be treated as a compound with deleted inner brackets. At the same time, however, the verbal prefix has syntactic properties that are characteristic of independent syntactic units, such as the possibility of movement into a position outside the VP and even outside a subordinate clause. Given these properties, É. Kiss analyzes the verbal prefix as an independent syntactic unit that is selected lexically by the verb. The second question that should be answered is whether it should be characterized as a phrase or a head since once again some facts suggest phrasal properties while other facts support head-like properties. Such contradictory evidence is resolved in this book by analyzing the verbal prefix as a

phrasal constituent consisting of a mere head. As such it is a projection that is both minimal and maximal and is capable of acting both as a phrase and as a head.

In Chapter 4 É. Kiss examines the semantic and syntactic properties of the preverbal focus constituent in Hungarian. Semantically focus is defined as the element that expresses exhaustive identification from among a set of alternatives. This exhaustive identification in Hungarian is the function of the immediately preverbal focus constituent. É. Kiss lists several semantic tests discussed in Szabolcsi (1981b) that show the exhaustive identification function of focusing.

The most conspicuous constraint on the flexible word order of the Hungarian sentence is the obligatory "focus V VM" word order in sentences containing a focus constituent. There have been several proposals in the literature to account for this reversal of the unmarked "VM V" word order. É. Kiss rejects on empirical grounds the analyses where the complementary distribution of the focus constituent and the VM is accounted for by assuming that they occupy the same preverbal position. Instead she adopts Brody's (1990) analysis of focus generating a focus projection, the specifier of which is obligatorily occupied by the focus constituent. She does, however, slightly modify Brody's original account and does not assume verb movement into the focus head. On the basis of empirical arguments she proposes that the FP is an alternative to the AspP. In other words, the VP is extended either into an AspP in neutral sentences or into an FP in sentences containing focus. In this way it automatically follows that the VM in sentences containing a focus can stand anywhere behind the verb since it stays in its base generated position in the verb phrase.

Wh-questions are also discussed in this chapter since *wh*-phrases are analyzed as obligatorily focused elements with a [+focus] feature. Two types of multiple questions are dealt with. One associated with a pair-list answer and the other one requiring a singular answer. It is claimed that the two types of questions involve different syntactic structures. In questions requiring a singular answer, one of the *wh*-phrases occupies Spec,FP while the other one remains in situ inside the VP. As mentioned earlier, *wh*-questions involving *wh*-movement to Spec,FP do not exhibit superiority condition effects since all arguments of the verb are at an equal distance from Spec,FP. *Wh*-questions that trigger a pair-list answer have a different structure. É. Kiss claims that in these sentences only one of the *wh*-phrases is an interrogative operator occupying the Spec,FP position while the other *wh*-phrase is a distributive quantifier occupying the position of distributive quantifiers above Spec,FP.

Chapter 5 deals with the leftmost position of the predicate phrase which is occupied by distributive quantifiers. Since in Hungarian the Nuclear Stress Rule assigns phrasal stress on the left edge of phrases, in the intonation phrase represented by the predicate each maximal projection is assigned phrasal stress. Therefore distributive phrases bear heavy stress. É. Kiss examines the set of quantifiers that can occur in quantifier position. This set can be divided into two groups. There is a set of quantifiers that can only appear in quantifier position. This includes universal quantifiers and phrases modified by the additive particle *is* which means 'also', or modified by *még... is* 'even'. She claims that they are restricted to the quantifier position because they inherently have the feature [+distributive]. The other set of quantifiers can occur in quantifier position but they can also appear in focus position, in topic position and postverbally. These include positive existential quantifiers and numeral phrases. É. Kiss assumes that distributive quantifiers occupy the specifier slot of a DistP projection which dominates either FP if the sentence contains a focus phrase or

AspP. The DistP projection—just like the TopP projection—can have more than one specifier. The scope principle, which says that an operator must c-command its scope, is observed in Hungarian in visible syntax, at surface structure.

Negation is discussed in Chapter 6. Negation is performed by the negative particle *nem*, which can appear either before the verb or before the focus or before the universal quantifier. Preverbal negation and negation preceding the focus exhibit parallel behavior. Several syntactic, phonological and semantic facts support the analysis that the VP negating particle and the FP negating particle both project a NegP and the negative particle sits in the Neg head of this projection. The adjacency of the VP negating particle and the verb is accounted for by assuming that in such a sentence no AspP is projected above the VP and thus NegP is not an extension of AspP but an alternative to it. Semantic considerations support this claim since negation does in fact neutralize aspect.

Negation of the universal quantifier, however, is of a different kind. It represents constituent negation with the negative particle *nem* adjoined to the quantified noun phrase. Hungarian exhibits the phenomenon known in the literature as Negative Concord, meaning that several instances of negative pronouns (*se*-phrases) can occur together with the negative particle, yet multiple negative pronouns do not yield multiple negation semantically. É. Kiss makes the following two assumptions: (1) a *se*-phrase has the feature [+negative,+distributive], (2) the functional heads Neg and Dist, instantiating these features, can fuse, projecting a joint DistNegP. The *se*-phrases occupy the specifier position of this joint DistNegP and this specifier position can be iterated.

From Chapter 7 on, the book deals with the inner structure of various types of verb complements. Chapter 7 discusses the structure of the Hungarian noun phrase starting by introducing the basic syntactic layers of the noun phrase. É. Kiss assumes that the NP kernel can be extended to a numeral phrase (NumP) which can host the plural marker. This can further be extended into a quantifier phrase (QP) and finally into a definite noun phrase (DP). NPs, NumPs, QPs and DPs have different distribution across sentence positions. One of the most challenging and interesting areas of Hungarian syntax is the structure of the possessive construction. In Hungarian the possessive relation is marked on the possessed noun with a suffix indicating possessiveness and also an agreement marker which agrees in person and number with the possessor. The structure of the possessive construction was first elaborated on by Szabolcsi (1981a; 1983; 1994) and in this book É. Kiss gives a detailed summary of this "traditional" or "standard" analysis of possessive constructions, discussing the merits and shortcomings of Szabolcsi's claims. In Hungarian the possessive relation can be expressed in three different ways. One construction involves a caseless possessor, another construction has a dative marked possessor internal to the extended projection of the possessed noun and there is a so called cleft construction where the dative marked possessor is moved out of the extended projection of the possessed noun. The relation between the possessor and the possession is identical in all three constructions and therefore it is commonly accepted that the three constructions should have the same underlying structure. É. Kiss gives several pieces of empirical evidence that Szabolcsi's claim that the dative marked possessor is derived from the nominative marked possessor cannot be maintained. Instead, for É. Kiss—following den Dikken (1999) and Bartos (1999)—the primary variant of the possessive construction is that containing the dative marked possessor. The caseless possessor is claimed to be in the specifier position of the DP projection.

The topic of Chapter 8 is the formal properties and syntactic structure of postpositions in Hungarian. É. Kiss assumes that postpositions are similar to Cases in Hungarian and that they are functional heads which extend the noun phrase into a prepositional phrase. This prepositional phrase is syntactically head initial which means that the prepositional head turns into a postposition only in the morphological component due to the fact that the P has a [+suffix] feature. This feature forces the obligatory adjacency of the noun phrase complement and the postposition. É. Kiss discusses the difference between postpositions taking a noun phrase complement and postpositions taking a pronominal complement. In the latter case, the PP must be extended into an AgrP and the agreement morphemes appearing on the postposition are identical to those that appear on the possessed noun in the case of a pronominal possessor. An interesting property of postpositions discussed in this chapter is that they can be assigned the feature [+verb modifier] if their complement has been extracted. The remnant PP exhibits all the properties of a verb modifier. In neutral sentences it immediately precedes the verb and acts as an aspectualizer.

Chapter 9 provides an analysis of three types of non-finite phrases in Hungarian: infinitival phrases, adverbial participial phrases and adjectival participial phrases. All three types of non-finite phrases can merge with the same types of operator phrases that can appear in finite clauses. A non-finite VP can be extended by AspP, NegP, FP, DistP and TopP. The movement possibilities of constituents are determined by exactly the same constraints as those valid in finite verb projections. However, non-finite clauses do not have a Case assigner that could assign nominative Case to the subject, therefore the subject of non-finite clauses is represented by PRO (except for inflected infinitival phrases discussed below). PRO can be controlled by the subject or the object of a matrix predicate. Inflected infinitives represent a rather unique property of Hungarian. The agreeing infinitive bears the same agreement marker that appears in possessive constructions on the possessed noun and the subject of the infinitive can be a Case-marked lexical noun phrase or a pronominal or *pro*. Partially basing her analysis on Tóth (2000a), a monograph devoted to inflected infinitives, É. Kiss provides an analysis that tries to account for the presence of the agreement marker that is sometimes obligatory and sometimes optional. She claims that the source of dative Case in inflected infinitives is the *-a/-e* suffix on the infinitive which is argued in section 7 to be a dative Case assigner.

Whereas in subject and object control constructions the infinitival verb has its own theta-role to assign and the whole infinitival phrase is the argument of the matrix verb, there are other constructions where the matrix verb and the infinitival verb form a complex predicate and they together assign a theta-role to the arguments. Verbs participating in this complex predicate formation are auxiliaries and semi-auxiliaries. The properties of these complex predicates have inspired a great amount of work, much of which is included in a monograph (É. Kiss – van Riemsdijk 2004). É. Kiss briefly discusses the two main structures that verbal complexes can have. One is the so-called straight verbal complex and the other one is the inverse verbal complex. In a straight verbal complex the surface order of the elements corresponds to the underlying order. In non-neutral sentences, however, when a focus or a negative particle precedes the finite verb, the non-finite elements of the verbal complex can appear in an inverse order. In the inverse order the verbal elements must be strictly adjacent. É. Kiss assumes that the inverse order is the result of cyclic incorporation of the infinitival

elements of the verbal complex. There are, however, opposing views, see for instance Koopman – Szabolcsi (2000).

If the extended VP is merged with the adverbial suffix *-va/-ve*, then the resulting non-finite phrase functions either as an adverbial of manner or time or it can also function as a secondary predicate. É. Kiss discusses the previous analyses of adverbial participle phrases that have been proposed in the literature (see Komlósy 1994; Laczkó 1995; Alberti 1998 and Tóth 2000b). The main debate concerns whether the fact that the subject argument is suppressed in a predicative adverbial participle indicates syntactic passivisation or a passive stem is already present in the lexicon. Although É. Kiss does not take an unambiguous stand on which analysis to adopt or support, she seems to favour Alberti's claim that, in Hungarian predicative adverbial phrases, a kind of passivisation takes place which prefers the patient. The suppression of the agent is a corollary of this patient preference.

The last chapter of the book presents the internal structure of subordinate clauses: the position of the subordinate clause in a matrix sentence, the properties of relative clauses and finally two interesting phenomena: long operator movement and the licensing of parasitic gaps. In Hungarian, subordinate clauses are associated with either a pronominal or a lexical head. This pronominal or lexical head plays two roles. First, it picks up the Case that is assigned to the subordinate clause by the matrix verb. Second, it can represent the embedded clause in those matrix operator positions where a clausal complement cannot appear. Such positions are Spec,AspP; Spec,DistP; Spec,FP and Spec,TopP. A further constraint on the position of an embedded clause is that a *that*-clause in Hungarian cannot be internal to a lexical projection. The relation between the pronominal or lexical head and the embedded clause coindexed with it has been a matter of debate in the literature. In this book É. Kiss mentions three alternative analyses: one proposed by Kenesei (1992), where the clause and the pronoun form an expletive-associate chain; one proposed by Lipták (1998), where the pronoun is generated in Spec,CP of the embedded clause; and finally an account given in É. Kiss (1987) where the pronoun and the clause constitute a complex noun phrase. In the section discussing long operator movement, É. Kiss examines how these three alternative analyses fare in view of the empirical facts and the reader gets the impression that Lipták's approach solves the largest part of the problems connected to long operator movement.

As the above discussion suggests, this book provides a detailed investigation of a wide range of phenomena in Hungarian syntax and gives possible accounts based on analyses that have been proposed in the literature and on previous work by É. Kiss herself. It is both an invaluable summary of the results of contemporary syntactic research on Hungarian and an original work in all senses of the word. I recommend this book to anyone who is interested in issues concerning the structure of Hungarian.

Ildikó Tóth

References

Alberti, Gábor 1998. On passivization in Hungarian. In: Casper de Groot – István Kenesei (eds): Approaches to Hungarian 6 (Papers from the Amsterdam conference), 103–22. JATEPress, Szeged.

Bartos, Huba 1999. Morfoszintaxis és interpretáció: a magyar inflexiós jelenségek szintaktikai háttere [Morphosyntax and interpretation: The syntactic background of Hungarian inflectional phenomena]. Doctoral dissertation, Eötvös Loránd University (ELTE), Budapest.

Brody, Michael 1990. Remarks on the order of elements in the Hungarian focus field. In: István Kenesei (ed.): Approaches to Hungarian 3 (Structures and arguments), 95–121. JATE, Szeged.

Dikken, Marcel den 1999. On the structural representation of possession and agreement: The case of (anti-)agreement in Hungarian possessed nominal phrases. In: István Kenesei (ed.): Crossing boundaries: Advances in the theory of Central and Eastern European languages, 137–78. John Benjamins, Amsterdam.

É. Kiss, Katalin 1987. Configurationality in Hungarian. Akadémiai Kiadó, Budapest.

É. Kiss, Katalin – Ferenc Kiefer (eds) 1994. The syntactic structure of Hungarian (Syntax and Semantics 27). Academic Press, New York.

É. Kiss, Katalin – Henk van Riemsdijk (eds) 2004. Verb clusters. A Study of Hungarian, German and Dutch. John Benjamins, Amsterdam & Philadelphia.

Kenesei, István 1992. Az alárendelt mondatok szerkezete [The structure of subordinate clauses]. In: Ferenc Kiefer (ed.): Strukturális magyar nyelvtan 1. Mondattan [A structural grammar of Hungarian, Vol. 1: Syntax], 531–713. Akadémiai Kiadó, Budapest.

Komlósy, András 1994. Complements and adjuncts. In: É. Kiss – Kiefer (1994, 91–178).

Koopman, Hilda – Anna Szabolcsi 2000. Verbal complexes. MIT Press, Cambridge MA.

Laczkó, Tibor 1995. On the status of *való* in adjectivalized constituents in noun phrases. In: István Kenesei (ed.): Approaches to Hungarian 5 (Levels and structures), 125–52. JATEPress, Szeged.

Lipták, Anikó 1998. A magyar fókuszemelések egy minimalista elemzése [A minimalist analysis of focus raising in Hungarian]. In: László Büky – Márta Maleczki (eds): A mai magyar nyelv leírásának újabb módszerei III [Recent methods in the description of contemporary Hungarian 3], 93–115. JATEPress, Szeged.

Szabolcsi, Anna 1981a. The possessive construction in Hungarian: A configurational category in a nonconfigurational language. In: Acta Linguistica Hungarica 31: 261–289.

Szabolcsi, Anna 1981b. The semantics of topic-focus articulation. In: Jan Groenendijk – Theo Janssen – Martin Stokhof (eds): Formal methods in the study of language, 513–41. Matematisch Centrum, Amsterdam.

Szabolcsi, Anna 1983. The possessor that ran away from home. In: The Linguistic Review 3: 216–89.

Szabolcsi, Anna 1994. The Noun Phrase. In: É. Kiss – Kiefer (1994, 179–274).

Tóth, Ildikó 2000a. Inflected infinitives in Hungarian. Doctoral dissertation, Tilburg University.

Tóth, Ildikó 2000b. *-va* and *-ván* participles in Hungarian. In: Gábor Alberti – István Kenesei (eds): Approaches to Hungarian 7 (Papers from the Pécs conference), 237–58. JATEPress, Szeged.

Jacqueline Guéron – Jacqueline Lecarme (eds): The syntax of time (Current Studies in Linguistics 37). MIT Press, Cambridge MA, 2004, 760 pp.

This volume contains the updated versions of talks presented at the *International Round Table on the Syntax of Tense and Aspect,* held at the Université Paris 7, in November 2000. The 23 articles in the book are between 14 and 44 pages. The contributors to the volume (in the order of the articles, which corresponds to the alphabetical order of the (first) authors) are Dorit Abusch, Mario Barra-Jover, Alexandra Cornilescu, Denis Delfitto, Hamida Demirdache, Myriam Uribe-Etxebarria, Yves D'hulst, Mürvet Enç, Nomi Erteschik-Schir, Tova Rapoport, Abdelkader Fassi Fehri, Alessandra Giorgi, Fabio Pianesi, Jacqueline Guéron, James Higginbotham, Michela Ippolito, Angelika Kratzer, Brenda Laca, Jacqueline Lecarme, Beth Levin, Malka Rappaport Hovav, David Pesetsky, Esther Torrego, Susan Rothstein, Philippe Schlenker, Carlota S. Smith, Tim Stowell, and Karen Zagona.

The editors, Guéron and Lecarme, survey the range of questions and problems discussed in the book in the 25 page long Introduction, and to some extent I will follow their classification. The range of the linguistic problems included in the volume is fairly large, and in many cases exceeds the domain of pure syntax. We can find morphological, semantic or pragmatic analyses in several articles. A few major research topics recur in a number of chapters: (i) models of tense construal (Reichenbachian and other approaches), (ii) dependency of tense on context, (iii) anchoring of tense, (iv) sequence of tense (SOT) phenomena and temporal ambiguity in embedded clauses or infinitival complements (double access reading, DAR), (v) problems in distinguishing aspect from Aktionsart, (vi) the characteristics of imperfective tenses, (vii) licensing of arguments (i.e., the relations between the argument structure of a sentence and its temporal properties), and, last but not least, the presence of tense in the nominal domain. Because of lack of space, rather than attempting to review the volume by grouping the linguistic problems into the topics mentioned above (which is what Guéron and Lecarme do in their Introduction), I will survey the content of the book by arranging the articles in four thematic classes as follows. (1) Telicity and argument structure (Erteschik-Schir and Rapoport, Levin and Rappaport-Hovav, Guéron, Kratzer, Rothstein, Higginbotham, Pesetsky and Torrego, Lecarme, and Cornilescu); (2) SOT and DAR phenomena (Enç, Schlenker, Zagona, Barra-Jover, and Abusch); (3) imperfective tense (Delfitto, Ippolito, Giorgi and Pianesi, and Fassi Fehri); (4) and finally a fourth group, which contains the remaining chapters, that investigate specific topics (Laca, D'hulst, Smith, Stowell, Demirdache and Uribe-Etxebarria). Of course, this is only a rough classification, and none of the articles can be considered to be as homogeneous as suggested by it.

The topics examined in the first group **"telicity and argument structure"** are quite heterogeneous: differences among Vendler-classes, problems of Aktionsart, tense chain in the nominal domain, and the analysis of structural cases by means of temporal features.

Cornilescu examines the Romanian infinitive (INF) and supine (SUP) nominalisations. She claims that while the behaviour of Romanian INF-nominals fit into the earlier analyses of Romance nominalisations by Kupferman and Alexiadou, the SUP-nominals contradict them. She concludes that it is necessary to make several modifications on Kupferman's and Alexiadou's theories. The author analyses two characteristics of Romanian SUP-nominals: (i) in the [nominal+subject] structures they behave like E(vent)-nominals, and not like R(esult)-nominals, (ii) they can license zero objects.

Cornilescu's explanation—based on Borer's (1994) theory—suggests that the differences between the Romanian INF- and SUP-nominals are due to an aspectual contrast between the two morphemes: the INF-suffix has a [+telic] feature, so it requires an overt object, while the SUP-suffix can have [+/−telic] features, as well.

Erteschik-Schir and Rapoport present a theory of structure projection of verbs, that determines thematic and aspectual interpretation, too. They suppose that syntactic structure is projected from the meaning components of the verb. They make use of three such meaning components ("bound semantic morphemes"): M (manner/means/instrument), S (state), L (location). In this analysis a verb is transitive just in case it has two such meaning components, and furthermore a verb with only one meaning component (e.g., *laugh*) can also become transitive when merged, for example, with a prepositional phrase. Following Hale and Keyser (1991) the authors claim that each pattern of the meaning components (and the syntactic structures derivable from these patterns) has a specific interpretation (e.g., activity, change-of-state achievement, etc.). A further characteristic of the theory is that theta-roles are not primitives, but derived.

Levin and Rappaport-Hovav's paper is closely related to Erteschik-Schir and Rapoport's analysis just mentioned, for Levin and Rappaport-Hovav also make the distinction among telic and atelic predicates on the basis of "event complexity" (they use this term). However, this concept cannot be unequivocally aligned with the above-mentioned "number of meaning components". Levin and Rappaport-Hovav think that the impact of traditionally recognised aspectual properties (e.g., telicity, boundedness) is overestimated, and they introduce the notion of event complexity instead. The Argument-per-Subevent Condition requires verbs expressing complex events to have objects. The authors claim that the classes of telic and complex events do not coincide. Complex events have the property that the authors call "lack of temporal dependence," i.e., the two subevents need not necessarily unfold together temporally. This is what we find, for example, in the case of reflexive resultatives and lexical causatives. For instance, the sentence *Sam has sung himself hoarse* is compatible with the following context: Sam sang yesterday enthusiastically during the class play, and when he woke up today, he was hoarse. By using this criterion the authors point out that the so-called verbs of consumption (e.g., *eat*), albeit being telic, are not complex events.

Rothstein examines two constructions in which accomplishment VPs are derived from nonaccomplishment heads: (i) progressive achievements, and (ii) resultative constructions with activity verbs. Rothstein assumes that in both cases type-shifting operations take place, since (i) achievement verbs express near-instantaneous changes-of-state (and so they are not compatible with the progressive) on the one hand, and, on the other hand, (ii) activity verbs do not have a culmination subevent (which is necessary for a resultative construction). In the author's view the possibility of transforming achievement and activity verbs to accomplishment predicates is based on the fact that the class of accomplishments bears similarities to both of the other classes: they extend over time (like activity verbs) and have inherently determined endpoints (like achievements). However, there are also differences between the derived accomplishments and lexical accomplishments: in the case of the former, the relation between verbal head and direct object is not gradual or incremental, but holistic.

Higginbotham modifies the treatments of Parsons (1990) and Landman (1992) in his analysis of the English progressive. Higginbotham's major innovation over Parsons and Landman consists in his representation of accomplishments and achievements, which directly and explicitly includes the telos part. While Parsons states that ac-

complishment verbs in English can only turn into a culminated event by adding the simple past tense morpheme to them, Higginbotham argues (following Zucchi 1999) that a verb actually can be an accomplishment verb even when the culmination is not attained. We can find a similar analysis in Erteschik-Schir and Rapoport's paper, who maintain that achievement and accomplishment verbs can actually be telic or atelic, as they may describe a single, final change of state, as well as the increments of that change.

Kratzer's article investigates the relation between telicity and the accusative case (ACC) in English and German. Following Pesetsky and Torrego's theory (see below), she assumes that verbal inflectional features might be the interpretable counterparts of uninterpretable case features, and that the relation between ACC case and telicity is agreement. Kratzer also makes use of a widespread analysis according to which events described by transitive verbs culminate with respect to the direct object referent. Direct objects in Finnish can bear two kinds of cases: ACC expresses the telicity of an event, while in case of atelic events the object has partitive (PART) case. The fact that in English and German ACC can appear also with atelic verbs complicates the picture. That is why Kratzer writes that German is like Finnish without PART-case. On the analogy of the morphologically overt imperfective (IMP) operator found in Russian, Kratzer suggests that in German there is a covert IMP-operator, and it occasionally neutralises the effect of the [telic] feature that is necessary to check ACC-case.

Guéron continues her earlier investigations in the area of tense construal. She posits that lexical items and grammatical morphemes both have [+/−ext(ended)] Aktionsart features. The value of the (spatial) Aktionsart depends on the inherent Aktionsart features of the lexical items inside the vP, and on the other hand it determines the value of the (temporal) aspect, i.e., a spatially extended ([+ext]) configuration in vP is construed as a temporally extended event at the level of Tense Phrase (TP). Furthermore, a sentence also needs a Tense Controller in [Spec,TP], which in the majority of cases is a subject with a [+human] feature. Its role is to license the temporal extension internal to an event, and at the same time the continuity between the temporality internal to, and the temporality external to, the event. According to Guéron, certain nouns (DPs) in the sentence possess an internal temporality (so-called biography), and in the time construal of the sentence the event the vP denotes is placed within the subject's time (biography). The subject's role is even more complicated: verbs with [+ext] Aktionsart feature select a "spatial subject", and its spatial contours delimit the configuration denoted by the vP. The object is also a kind of measure: it delimits the number of gestures the subject needs to perform in order to achieve the spatial configuration. In Guéron's opinion achievement verbs, for instance, have [−ext] features, while accomplishments are of [+ext]. Besides, auxiliaries in this theory also have content (as opposed to Chomsky's analysis), and this content is exactly a [+/−ext] feature. When an auxiliary raises to Tense (T), its Aktionsart feature combines with T and is construed as imperfective or punctual aspect.

There are two articles left in the group referred to here as "telicity and argument structure": one by Pesetsky and Torrego, and one by Lecarme, which is to a very large extent based on their theory.

Pesetsky and Torrego further develop their earlier theory in Pesetsky and Torrego (2001). Relying on Chomsky (1995), they argued there that an essential ingredient (and trigger) of movement is an Agree relation between an uninterpretable feature (uF) of a so-called probe category and a corresponding feature of a so-called goal

category. Nominative case (NOM) is considered to be an instance of an uninterpretable Tense feature (uT) on the head of a DP category, and subject agreement on the verb reflects uninterpretable ϕ-features ($u\phi$) on Tense (T). Furthermore, there is an Agree relation between the uT on D and $u\phi$ on T. In their present article the authors extend this analysis, claiming that all instances of structural case are instances of uT on D. Pesetsky and Torrego investigate several phenomena on the basis of this hypothesis, of which I will only sketch two: (1) the *that*-trace effect, and (2) the accusative (ACC) case. In the case of the first of these, the authors suggest that the word *that* in English is actually not a complementizer, but a realization of T moved to C (i.e., the Complementizer position). English C is then phonologically null, and may trigger either T-to-C movement (with the word *that*) or subject movement to Spec, CP (without *that*). The treatment of ACC is analogous to NOM: in the case of verbal predicates the authors stipulate the presence of a second occurrence of T, which they label $T_{O(=object)}$. The fact that adjectival and nominal predicates cannot have DP complements in English is not given a unified answer. In their opinion, in the case of adjectival predicates T_O is absolutely absent, but nominals do have a defective T_O (nominal T_O requires complements with interpretable T-features, e.g., PPs). Pesetsky and Torrego claim that this is because nominals lack a full tense system.

Lecarme examines the Tense features in the Somali DPs. She uses the theory of Pesetsky and Torrego (2001) as theoretical background, but goes beyond it in two important respects. First, according to her analysis in Somali the tense/event structures of nominals and clauses do not differ, i.e., the chain Op_i, T_i, e_i is available in both of them (C-T-V and D-T-N). Morphological parallels among the verbal and the nominal domain provide evidence. Second, structural genitive is a reflex of the D-T relation, analogously to the NOM, which is the reflex of the C-T chain according to Pesetsky and Torrego (2001).

I begin the introduction of the group of articles labelled **SOT and DAR phenomena** with the chapter by **Enç**. Enç finds it necessary to revise her earlier analysis (Enç 1987), according to which all occurrences of surface past tense behave like a past tense. This time the author extends her so-called Anchoring Conditions from T-nodes to inflectional nodes (I): assuming that each I must be temporally anchored, and that each I carries two temporal indices (evaluation index and referential index). An I is temporally anchored if (i) it is bound by a local c-commanding I, (ii) its evaluation time is bound by a local c-commanding I, or if (iii) its evaluation time is fixed as the speech time. Only I's with the feature [+past] can bind another I. If a past tense is embedded under a future-shifting modal, it shifts back from the future time. Furthermore, if a nonfinite I is embedded under an I having [+past] feature, the former inherits the feature [+past], and in this way it becomes capable to bind another I.

Barra-Jover's article deals with the differences between the syntax of French and English direct quotations. He points out that the introductory statements behave differently in sentence-initial and in noninitial positions. According to the Anchoring Conditions for tense established by Enç (1987), if a C does not have a governing category, it is anchored if and only if it denotes the speech time. Barra-Jover weakens this Anchoring Condition by claiming that if C does not have a governing category, it is anchored just in case it denotes the speech time or if there exists an accessible antecedent denoting the speech time. Besides, the author states that in simplex sentences tense (T) is specified in case it is [+E] and [+S] at the same time, i.e., it refers to the event time and to the speech time, too. He distinguishes between two sorts of underspecified

tense: (1) subjunctive is [−E][−S], while (2) quotations are [+E][−S], i.e., the speech situation of the sentence cannot be identified with that of the speaker. The weakened version of the Anchoring Condition can be used in the latter case: in direct quotations it is necessary to postulate an operator (OpT) in C having an accessible antecedent. The antecedent can be either lexical (i.e., initial introductory statement), or nonlexical (time of the preceding series of narrative events).

Schlenker analyses SOT and DAR phenomena not only in the temporal domain, but in connection to pronouns and moods, too. According to him, the semantics of pronouns, tenses and moods are similar: pronominal, temporal and modal features are presuppositions on the values of individual, time and world variables, respectively. The sequence of tense (SOT) means that tense features of a T_2 embedded under a T_1 are eliminated, and T_2 inherits the tense features of the upper T_1. Similar phenomena can be encountered in the two other domains, too: for instance, when the [+masculine] feature of an embedded pronoun is deleted, or when the indicative mood features of an embedded predicate are ignored in the interpretation. Schlenker treats all these phenomena in a unified fashion: by stipulating purely morphological rules of agreement. He borrows the idea from Heim (1994), who observed that in some cases a pronoun embedded under an attitude verb cannot be literally interpreted as coreferential with an argument of the superordinate clause, and suggested that there are only purely morphological rules of agreement at work. Schlenker generalises the idea by claiming that context variables embedded under attitude verbs inherit the features of the individual, time, and world arguments of the embedding verb.

Zagona's article deals with double access readings (DAR) in Italian and Spanish. She follows Giorgi and Pianesi's (2000) analysis in assuming two complementizer positions in Italian: an upper (standard) C, and a lower one. Zagona states that verbs of communicative behavior select the upper C, and as a result they never display complementizer deletion. Attitude verbs in Italian, in turn, select the lower C, resulting in the possibility of complementizer deletion and this blocks DAR. The author claims that in the background of the differences between the two matrix verb classes there is an aspectual contrast: verbs of communicative behavior are processes, while verbs of pure attitude are states. Zagona draws a comparison between verbs of attitude, and adjectival predicates selecting a CP-complement: both classes are stative predicates, and neither of them shows the DAR. Besides, we can also see an important difference between Zagona's and Giorgi and Pianesi's points of view: according to the latter, a T in AgrP position is anchored to the matrix event, while a T raised to C is anchored to the speech time. In contrast, in Zagona's opinion the T raised to C is anchored to the matrix event, and T-to-C movement is triggered by the aspectual properties of the matrix verb: nonstative matrix verbs (e.g., verbs of communicative behavior) select the upper C, which in turn triggers T-to-C movement in the embedded clause.

The last among articles on SOT and DAR phenomena to be discussed here is by **Abusch**. It investigates the logical form of English *to*-complements. Using the presence (or lack) of simultaneous and future oriented interpretations in *to*-complements as a criterion, Abusch classifies verbs selecting *to*-complements in three groups. (1) The so-called B-verbs (e.g., *believe*) permit only simultaneous interpretation in their complement, while (2) the so-called F-verbs (e.g., *forecast*) license both simultaneous and future oriented interpretations of their complement. (3) Besides, there exist a few verbs the *to*-complements of which can obtain only futurate readings (e.g., *hope*). In her analysis of future-oriented infinitives Abusch uses the semantics of the future aux-

iliary *will* as a model: in her view, the core meaning of *will* is a temporal substitution operator, locating the eventualities corresponding to the main verb in the interval (t, ∞), i.e., positive infinity, where t is a time variable. The author assumes that the representation of future oriented infinitives does also contain this temporal substitution operator, and in case of sentences permiting both simultaneous and futurate readings, she uses the interval $[t, \infty)$, including the left boundary t.

We find four articles in the book dealing primarily with **imperfective tense**. Three of them investigate the Italian imperfective tense, while the fourth one sets out to answer the question of whether or not Arabic is an "aspect language".

Delfitto argues that imperfective tenses (IMP) are uniformly mapped into subject-predicate logical formats, and in Germanic and Romance languages grammatical aspect is the locus where the distinction between categorial and thetical sentences is grammatically encoded. Although traces of displaced arguments are usually not interpreted as predicational traces, this is exactly what the IMP marking is supposed to do according to Delfitto: IMP tenses encode the information that one of the verb's arguments has to be interpreted predicationally. Technically speaking, in the case of IMP marked verbs the author stipulates a functional projection PredP, and assumes that one of the arguments has to be displaced to Spec, PredP. In Romance many of the sentences involving left-dislocated topics receive a thetical interpretation: in these cases topics undergo VP-internal "logical" reconstruction. The role of IMP is to encode that one of the verb's arguments is not allowed to reconstruct. So while perfective (PERF) marking indicates that the VP is viewed as a fully saturated expression, the IMP marking refers to subject-predicate logical format. The subject of predication does not always coincide with the grammatical subject: for instance, in the case of the progressive reading of IMP the logical subject is the evaluation time t.

Giorgi and Pianesi follow earlier work by Delfitto and Bertinetto (2000) in treating TP as an argument of the verb, and they also acknowledge the generalization that IMP needs a temporal topic. The authors argue that tenses behave as shiftable indexicals, i.e., in many cases they refer to the temporal coordinates of the attitude's subject rather than picking out that of the speaker's. The temporal topic of an embedded clause is identified with the event time of the matrix clause, and there are also cases (e.g., dream contexts) in which anchoring is not enforced. Following others, Giorgi and Pianesi assume that if a clause expresses the content of a propositional attitude of a subject, then its "interpreted logical form" contains his/her egocentric coordinates. In embedded clauses the attitude by a subject is often distinct from that of the speaker's. Furthermore, IMP is not evaluated with respect to the speaker's actual coordinates in matrix clauses either: instead, it expresses past expectations concerning a tenseless proposition.

The third article in the volume on the Italian IMP is **Ippolito's**. She investigates modal and conditional uses of the imperfect. Since in modal uses of the IMP (in contrast to the aspectual readings) the relevant eventualities are not necessarily understood as past, Ippolito argues that in such cases the function of IMP is the restriction of an accessibility relation, rather than locating the event in the past. This accessibility relation is a binary relation between a world-time pair and a set of worlds compatible with it. A sentence with the IMP is true just in case the proposition is true in all the worlds that were accessible to the speaker in the actual world at a time prior to the utterance time. This is what Giorgi and Pianesi call past expectations of the speaker. Ippolito also analyses the uses of the Italian IMP in conditionals in detail. She

points out that while in subjunctive conditionals the implicature that the antecedent is false seems to be in general cancellable, in case of IMP conditionals this implicature is noncancellable. The author suggests that the meaning that "the speaker believes that not-p" found in IMP conditionals can be derived by scalar implicature from the more fundamental meaning "the speaker does not know that p" in modal uses of IMP.

Fassi Fehri's analysis of Arabic can be placed in the context of the longstanding debate on whether Semitic are "aspect languages", i.e., whether the category of tense is really absent in them or not. In Arabic the same inflected verbal form can express Past (Non-Past) and Perfective (Imperfective) senses, i.e., there is no morphological distinction between Agr_1 and Agr_2, or between T_1 and T_2. In spite of these facts, the author denies the existence of verbal Aspect as a (discrete) grammatical category in Arabic and considers the Past/Present tense opposition to be the primary function of the relevant morphological tools. This analysis contradicts the traditional Western grammars on Arabic, but harmonises with the standpoint of the traditional Arabic grammarians, and with that of Kuryłowicz (1972).

The articles in the **fourth group** are those whose topics do not really fit in the previous three classes, but this does not mean that the five articles left do not have many connections to those surveyed previously.

Demirdache and Uribe-Etxebarria integrate time adverbs into their earlier analysis of Tense and Aspect. In their earlier work tenses and aspects are dyadic predicates of spatiotemporal ordering. (Giorgi and Pianesi 1997 have a similar analysis, replacing the Reichenbachian ternary relation of reference time (R), speech time (S) and event time (E) by two binary relations: S-R (T_1), R-E (T_2).) Time adverbs are considered to be semantically and syntactically restrictive modifiers of the temporal arguments projected by Tense and Aspect. Tense and Aspect as being dyadic predicates relate their temporal arguments on the basis of a unique basic semantic opposition: [+/−central coincidence] between the location of the figure and that of the ground (cf. Hale 1984). The feature [+centr.co.] expresses F(igure) WITHIN G(round), while [−centr.co.] expresses F BEFORE/AFTER G. Time adverb phrases are also headed by a two-place predicate of spatiotemporal ordering. This head can be either overt in syntax (PP adverbs) or covert (e.g., *last year*). In the sense of this analysis temporal adjunct clauses (*when*-clauses) are also headed by a covert preposition, and silent prepositions always express [+centr.co.].

D'hulst investigates the historical development of synthetic conditional tenses in Western Romance on the basis of Roberts' (1992) view on the grammaticalization process of Romance synthetic futures. In most of the Romance languages conditional morphology is based on the Vulgar Latin periphrastic construction using imperfect morphology on *habere* 'have'. However, in Italian the actual conditional has developed from the Vulgar Latin forms with perfect tense on *habere*. The author gives the following explanation for the origin of 'future-in-the-past' meaning in the case of the conditional forms mentioned: the originally biclausal periphrastic structure changed into a monoclausal construction by the embedded verb climbing up to the matrix clause, ending in the reanalysis of lexical *habere* as an auxiliary. Therefore in the new picture a past tense auxiliary dominates an infinitive expressing future, and this is exactly what is required in order to express future in the past. As a result of this process, the future on the INF was reanalysed as T_2 of the earlier matrix clause. Still later the future value of T_2 shifted to T_1, opening the way to the development of composed future forms (having participle forms in T_2).

Laca also examines the Romance languages, but from the point of view of another problem: she investigates the so-called aspectual periphrases (e.g., French *aller à* + INF 'to be going to + V'). She disagrees with the treatment of Cinque (1999), who proposed that the higher/finite verbs in these constructions are "functional" verbs, because this approach ends in a proliferation of functional heads. Instead, Laca claims that such finite verbs distribute over two levels of structure: (i) a lower level containing verbs encoding Aktionsart, and (ii) an upper level with verbs encoding syntactic aspect.

Stowell poses the question of whether English modals (e.g., *could, might*) should really be considered to involve a morphosyntactic combination of tense with a modal verbal head, as the present/past alternation is semantically neutralised for these verbs in many contexts. (For instance, Enç in this volume states that in the case of *would, could, might*, etc. past shows up on the morpheme for historical reasons.) Modal verbs can in general have two readings: in the epistemic uses they may not fall under the logical scope of tenses, while in the root modal uses they are free to do so. The evaluation time of *might, ought* can be in the past, when they are governed by an intensional verb in a past tense main clause (SOT). However, in the same environment *may* and *must* require a double access reading (DAR). Stowell states that this contrast suggests that epistemic modals like *might, should, ought to* do in fact involve an occurrence of the morphological past tense.

The last article to discuss in this review is **Smith**'s on tense interpretation in various genres of discourse (or in her term discourse modes). She investigates five discourse modes (Narrative, Description, Argument, Report, and Information) in the theoretic background of DRT. The two questions she posits about each discourse mode are: (i) what type of entity is introduced into the universe of discourse, and (ii) what principle of advancement organises tense interpretation? There is a strong correlation to be found between the discourse modes and the types of entities (e.g., eventualities, generic statements, propositions, etc.) in them. Advancement in so-called temporally organised modes (i.e., Narrative, Description, Report) takes place with respect to location changes in time or space, whereas in case of atemporal modes (i.e., Argument and Information) with respect to metaphorical location and motion. In the case of the Narrative the author follows directly the analysis of Kamp and Reyle (1993), but in other cases (e.g., Description) she modifies it.

This review could only sketch part of the problems and ideas that can be read about in the volume. An essential merit of the book *The syntax of time* is that it provides not only an exhaustive enumeration of phenomena currently investigated in the domain of tense and aspect, but it also presents a number of theories as well. The linguistic data covered in the articles are also of considerable richness: a total of 27 languages from several language families are listed in the integrated Index of languages, subjects, and authors, a customary feature of books published by the MIT Press, though in this case the Index could have been more carefully compiled, since some of the crucial terms discussed in important papers do not figure, such as *achievement* and *accomplishment* in Higginbotham's article. Another thing missing in the book is a unified bibliography, as references are listed at the end of the individual articles. But these minor shortcomings hardly diminish the overall value of this excellent book.

Balázs Szilárd

References

Borer, Hagit 1994. On the projection of arguments. In: Elena E. Benedicto – Jeffrey T. Runner (eds): Functional projections (University of Massachusetts Occasional Papers in Linguistics 17), 19–47. University of Massachusetts, GLSA, Amherst.

Chomsky, Noam 1995. The minimalist program. MIT Press, Cambridge MA.

Cinque, Guglielmo 1999. Adverbs and functional heads. Oxford University Press, New York.

Delfitto, Denis – Pier Marco Bertinetto 2000. Word order and quantification over times. In: James Higginbotham – Fabio Pianesi – Achille Varzi (eds): Speaking of events, 207–43. Oxford University Press, Oxford.

Enç, Mürvet 1987. Anchoring conditions for tense. In: Linguistic Inquiry 18 : 633–657.

Giorgi, Alessandra – Fabio Pianesi 1997. Tense and aspect: From semantics to morphosyntax. Oxford University Press, Oxford.

Giorgi, Alessandra – Fabio Pianesi 2000. Sequence of tense phenomena in Italian. In: Probus 12 : 1–32.

Hale, Kenneth 1984. Notes on world view and semantic categories: Some Walpiri examples. In: Pieter Muysken – Henk van Riemsdijk (eds): Features and projections, 233–54. Foris, Dordrecht.

Hale, Kenneth – Samual Jay Keyser 1991. On the syntax of argument structure. Lexicon project working paper 34. MIT Press, Cambridge MA.

Heim, Irene 1994. Puzzling reflexive pronouns in de se reports. Handout, Bielefeld Conference, March 1994.

Kamp, Hans – Uwe Reyle 1993. From discourse to logic. Kluwer, Dordrecht.

Kuryłowicz, Jerzy 1972. Studies in Semitic grammar and metrics. Curzon Press, London.

Landman, Fred 1992. The progressive. In: Natural Language Semantics 1 : 1–32.

Parsons, Terence 1990. Events in the semantics of English: A study in subatomic semantics. MIT Press, Cambridge MA.

Pesetsky, David – Esther Torrego 2001. T-to-C movement: Causes and consequences. In: Michael Kenstowicz (ed.): Ken Hale: A life in language (Current Studies in Linguistics 36), 355–426. MIT Press, Cambridge MA.

Roberts, Ian 1992. A formal account of grammaticalisation in the history of Romance futures. In: Folia Linguistica Historica 13 : 219–58.

Zucchi, Alessandro 1999. Incomplete events, intensionality, and imperfective aspect. In: Natural Language Semantics 7 : 179–215.

Ken Safir: The syntax of (in)dependence (Linguistic Inquiry Monograph 44). MIT Press, Cambridge MA 2002, xvi + 194 pp.

Ken Safir's aim is to reformulate one of the fundamental building blocks of post-*Aspect* generative grammar, Binding Theory, in simpler and more general terms. His ambitious project is to be realized in three instalments, the first two of which have already been published as Safir (2004) and the book under review here, with the third

one in the making. The book is divided into five chapters and an Appendix. Chapter One is an introduction to Safir's objectives and an overview of the relevant precursors including Safir (2004). He proposes to fuse as well as supersede Binding Principles B and C, which regulate the coreferential properties of pronominals and r-expressions (or names) by his Form to Interpretation Principle (FTIP), which has the form in (1), as based on the competing principles in (2) and (3):

(1) If x c-commands y and z is not the most dependent form in position y with respect to x, then y cannot be directly dependent on x.

(2) INDEPENDENCE PRINCIPLE (INP)
If x depends on y, then x cannot c-command y.

(3) C-COMMAND LICENSING PRINCIPLE (CLP)
If x depends on y, then y must c-command x.

The FTIP is applied relative to a "Most dependent hierarchy", in which anaphors are more dependent than pronominals, which in turn are more dependent that names (r-expressions). Incidentally, this picture of gradual dependency puts Safir's proposal in the neighborhood of Optimality Theory, where such a progression is more easily coped with than in the Principles and Parameters Theory of Chomskyan (mainstream?) generative grammar. The INP and the CLP differ in the domains of their applications, as transpires from Chapter Two, "The distribution of dependency", which lists examples and arguments in favor of the INP, and endorses Fox's (1998) Rule H, which acts as a locality constraint in a c-command chain, in which the closest c-commanding item is taken as an antecedent. Chapter Three bears the title "Deriving crossover": it offers a unified treatment of crossover phenomena as subsumed under the INP with an extension to include dependency relations relative to quantifiers, called Quantifier Dependency Condition (QDC). Chapter Four, "Reconstruction and dependent readings", argues for the copy theory of movement. It is here that (in)famous examples containing so-called "picture-nominals" (e.g., *Which picture of **Bill** was **he** afraid that Hillary would be thinking of?*) and problems of late adjunction are discussed at length (i.e., the difference between complement clauses and relative clauses with respect to the coreferential behavior of pronominals in them). The last chapter, "The Independence Principle in the architecture of Universal Grammar", concludes that the place where the relevant principles are at work is the Logical Form, and since it is an interpretive mechanism, it has to make a case against a movement analysis of the interpretation of dependent nominals, as proposed by Hornstein (2001) and Kayne (2002), supported by crucial examples from weak crossover and the necessity for both to make reference to Binding Principle A. Recall that Safir makes do without the Binding Principle, since his INP, FTIP, Rule H, and QDC are sufficient to cover all cases of quantifier-pronominal interpretation as well as familiar cases of binding. The ten-page Appendix is an attempt to accommodate data from Hindi/Urdu that shows extensive scrambling.

Along the way a number of side issues receive adequate treatment or at least some attention, such as the problem of "proxy terms", as the pronoun in ***Marlene** thought **her** nose was too long*, where bolded items are "coconstrued"—to use Safir's term, or the intriguing behavior of PRO in weak crossover, cf. ***Who** did* [**PRO** *shaving **himself***] *convince **t** to grow a beard?* vs. * ***Who** did* [**PRO** *shaving **himself***] *convince Mary to*

trust t?, which show that PRO is immune to weak crossover effects under conditions similar to quantifier-pronominal dependencies, or "vehicle change", as in *Mary loves **John** and **John** admits she does (love **him**/***John**), too*, in which the name has to give way to the pronoun in the grammatical version underlying the elliptical form.

The book is well-organized and well-argued. Although it is not always an easy read, for some of the arguments rely on quite complex data, almost exclusively from English (discounting the Appendix), it is a must for anyone interested in the thorny problems of binding and quantifier-variable interpretation and in the debate on whether binding phenomena can be handled by means of a Probe-and-Agree type analysis as suggested by recent developments in the Minimalist Program or by a set of general enough principles ultimately (also) based on lexical characterizations. As is usual with the MIT Press, the book is neatly produced, although I missed some of the terms in the Index, e.g., the DSV (=definition of syntactic variable) approach, or QDC (=Quantifier Dependency Condition). The only typo I have noticed is hardly of significance: the page numbers of Chomsky and Lasnik (1995) in the References fit the article that follows it in the book, *The minimalist papers.*

István Kenesei

References

Fox, Danny 1998. Locality in variable binding. In: Pilar Barbosa – Danny Fox – Paul Hagstrom – Martha McGinnis – David Pesetsky (eds): Is the best good enough? Optimality and competition in syntax, 129–55. MIT Press, Cambridge MA.

Hornstein, Norbert 2001. Move! A minimalist theory of construal. Blackwell, Malden MA & Oxford.

Kayne, Richard S. 2002. Pronouns and their antecedents. In: Samuel David Epstein – Daniel T. Seely (eds): Derivation and explanation in the Minimalist Program, 133–66. Blackwell, Malden MA & Oxford.

Safir, Ken 2004. The syntax of anaphora. Oxford University Press, Oxford.

Acta Linguistica Hungarica, Vol. 53 (1), pp. 95–96 (2006)

HUNGARIAN BOOKS ON LINGUISTICS

László Kálmán – Viktor Trón: Bevezetés a nyelvtudományba [An introduction to linguistics]. Tinta Könyvkiadó, Budapest, 2005. 155 pp.

This volume summarizes the material covered by introductory linguistics courses as held by the authors for several years. Therefore, it can be used as a companion to such courses in higher education, but also as an introductory reading for the interested layman. It is certainly a practical aid for teaching linguistics courses although it does not contain exercises.

It is true that books with this title abound in many languages; this particular one, however, has a couple of exceptional features. First, it is the first one in Hungarian in the last 30 years or so. Since many concepts and terms used in current linguistics have been introduced or redefined in the past few decades, the book will be indispensable in teaching introductory linguistics courses in higher education. Second, it keeps a distance from contemporary linguistic theories. It presents most concepts used in mainstream theories as problematic rather than self-evident, but it does not take a stance on their status as a rule. The authors' sceptic view on most common distinctions (such as the distinction between arguments and adjuncts or inflection and derivation) gives the book a more critical flavour than usual with introductory coursebooks.

The somewhat unusual structure of the book may also originate in the authors' basic stance. The introductory chapter (on the concept of language and linguistics) is followed by one about the cognitive abilities of humans and their relationship to language, a topic that is usually deferred in such manuals. The next chapter is about varieties of language, i.e., the sociological aspects of language use, which is also among the last ones in standard textbooks. The traditional sub-disciplines of linguistics are also treated in a somewhat unexpected way. The chapter on 'the use of language', which covers both semantics and pragmatics, without any clear-cut boundary between the two, precedes rather than follows those on phonetics and phonology (again, without a clear-cut borderline), morphology and syntax. The central role that syntax plays in

1216–8076/$ 20.00

many contemporary theories is missing in terms of both the position of the chapter within the book and its length. Finally, a chapter on historical linguistics and a lengthy study on the history of linguistics follow.

The book is well-organized on the whole, although it is rather compact in its style and sometimes parsimonious in terms of examples. The readers can also benefit from the extra material at the end of the book: an up-to-date collection of background readings, a glossary of all important concepts mentioned in the book as well as a detailed index.

INFORMATION FOR CONTRIBUTORS

Please provide typewritten double-spaced manuscripts with wide margins (50 characters per line, 25 lines per page). Type on one side of the paper only and number all the pages consecutively. The following information should be provided on the first page: the title of the article (or, in case of reviews, the title of the book under discussion, the name of the publisher, the place and date of publication as well as the number of pages), the author's full name and address (including email address) and an abbreviated title of the paper (not exceeding 45 characters including spaces). Only original papers will be published and a copy of the *Publishing Agreement* will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement. Authors are requested to send two hard copies of their manuscript + a floppy disk under DOS/Windows. The manuscript should be accompanied by a list of five keywords and an abstract of about 100 words. It should be typed on a separate sheet of paper. Tables, diagrams and illustrations (with the author's name and an Arabic number) should be presented on separate sheets. Captions should be typed on a separate sheet and placed at the end of the manuscript. Footnotes should be typed double-spaced on a separate sheet at the end of the article. All language examples (and only these) are to be italicized (single underlining). Citations in the text should be enclosed in double quotation marks (“ ” in case of a paper written in English, „ “ in German and « » in French).

Reference to a publication should be made by the name of the author, the year of publication and, when necessary, by page numbers in the following ways:

...as described by Schmidt (1967)...
...as referred to by Hoover (1967, 56–78; 1976, 43)...
...mentioned by several authors (Elgar 1978, 77; Williams 1981, 154–6).

An alphabetically arranged list of references should be presented at the end of the article as follows:

Bárczi, Géza 1958a. Magyar hangtörténet [The historical phonology of Hungarian]. Akadémiai Kiadó, Budapest.

Bárczi, Géza 1958b. A szótövek [Word stems]. Akadémiai Kiadó, Budapest.

Brockhaus, Wiebke 1995. Skeletal and suprasegmental structure within Government Phonology. In: Jaques Durand – Francis Katamba (eds): Frontiers in phonology: Atoms, structures, derivations. 180–221. Longman, Harlow.

Cole, Jennifer 1995. The cycle in phonology. In: Goldsmith (1995: 206–44).

Goldsmith, John A. (ed.) 1995. The handbook of phonological theory. Blackwell, Cambridge MA & Oxford.

Kaye, Jonathan – Jean Lowenstamm – Jean-Roger Vergnaud 1990. Constituent structure and government in phonology. In: Phonology 7: 301–30.

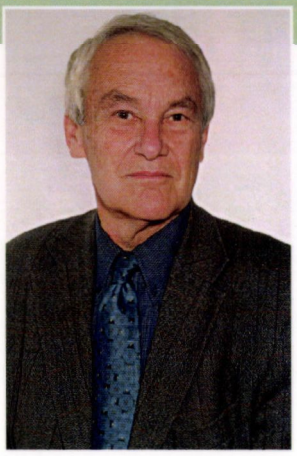
Tomioka, Satoshi 1997. Focusing effects and NP-interpretation in VP-ellipsis. Doctoral dissertation, University of Massachusetts, Amherst.

For marking subsections decimal notation should be applied. Do not use more than four digits if possible. Examples within the text should be marked in italics. Meanings are to be rendered between inverted commas (‘ ’). If glosses are given morpheme by morpheme, the initial letter of the gloss should be placed exactly below that of the example. Grammatical morphemes can be abbreviated in small case letters connected to the stem or the other morphemes by a hyphen. No period should be applied in such abbreviation. For example:

(1) (a) A sólymaid elszálltak
the falcon-gen-pl-2sg away-flew-3pl
‘Your falcons have flown away.’

Examples can be referred to in the text as (1a), (1a–d), etc.

One proof will be sent to the author. Please read it carefully and return it by air mail to the editor within one week.



FERENC KIEFER, editor-in-chief, professor emeritus
Member of the Hungarian Academy of Sciences
President of CIPL
Corresponding member of the Austrian Academy of Sciences
Member of Academia Europaea and of
European Academy of Arts, Sciences and Humanities
Main research areas:
morphology, lexical semantics, semantics-pragmatics interface
He has published 11 books (in Hungarian, English, French and German) and more than 200 articles

CODEN: ACTA LINGUIST. HUNG.
SIGILLUM: ALHUE 8

ISSN 1216-8076

phonology ■ morphology ■ syntax ■ semantics and pragmatics ■ sociolinguistics ■ cognitive linguistics

Volume 53 ▪ Number 2 ▪ June

2006

Editor-in-Chief ▪ FERENC KIEFER

FOUNDED IN 1951

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

WWW.AKADEMIAI.COM

AKADÉMIAI KIADÓ

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Acta Linguistica Hungarica publishes papers on general linguistics with particular emphasis on discussions of theoretical issuses concerning Hungarian and other Finno-Ugric languages. Papers presenting empirical material must have strong theoretical implications.
The scope of the journal is not restricted to the traditional research areas of linguistics (phonology, syntax and semantics, both synchronic and diachronic), it also covers other areas such as socio- and psycholinguistics, neurolinguistics, discourse analysis, the philosophy of language, language typology, and formal semantics.
The articles should be written in English, exceptionally, however, papers written in German and French are also accepted. The journal is published in yearly volumes of four issues.

■

Abstracted/indexed in
Bibliographie Linguistique/Linguistic Bibliography, International Bibliographies IBZ and IBR, Linguistics Abstracts, Linguistics and Language Behaviour Abstracts, MLA International Bibliography.

■

Editorial correspondence
ACTA LINGUISTICA HUNGARICA
Research Institute for Linguistics
P.O. Box 701/518, H-1399 Budapest, Hungary
Phone: (+36 1) 351 0413; Fax: (+36 1) 322 9297
E-mail: kiefer@nytud.hu; siptar@nytud.hu

■

Subscription price
for Volume 53 (2006) in 4 issues EUR 252 + VAT (for North America: USD 316) including online access and normal postage; airmail delivery EUR 20 (USD 25).

■

Please send your order to
AKADÉMIAI KIADÓ
Journals Business Centre
P.O. Box 245, H-1519 Budapest, Hungary
Phone: (+36 1) 464 8222; Fax: (+36 1) 464 8221
E-mail: journals@akkrt.hu
www.akademiai.com; www.akademiaikiado.hu

■

ISSN 1216-8076

ALing 53 (2006) 2

Printed in Hungary

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Volume 53, Number 2, June 2006

AKADÉMIAI KIADÓ
MEMBER OF WOLTERS KLUWER GROUP

Cover design: xfer grafikai műhely

CONTENTS

Acta Linguistica Hungarica, Vol. 53 (2), pp. 97–116 (2006)
DOI: 10.1556/ALing.53.2006.2.1

ANALYSIS OF THEORETICAL RESEARCH ARTICLE INTRODUCTIONS WRITTEN BY UNDERGRADUATE STUDENTS: A GENRE-BASED APPROACH

RÉKA FUTÁSZ

Center for Academic Writing
Central European University
Nádor utca 9.
H–1056 Budapest
Hungary
rfutasz@freemail.hu

Abstract: The present study investigates the rhetorical structure of theoretical research article introductions written by undergraduate students in the field of English linguistics and literature. The aim of the study was to see if English-language introductions written by Hungarian students majoring in English conformed to those typically written by native speakers, and to examine if two disciplines, English literature and linguistics observe different conventions in terms of rhetorical structure in introductions. The findings suggest that although on the whole the Hungarian introductions displayed the elements typically identified in research articles written by native expert writers, discipline-specific variations emerged.

Keywords: genre analysis, research article, linguistics, second language, student writing

1. Introduction

The typical structures of various genres have been studied extensively in recent years, especially following the work of Swales (1981; 1984; 1990). There are many situations where such analyses might help writers and readers in making sense of texts, though the vast majority of the research done in this field has focused on a genre which is typically important for those who conduct research themselves, i.e., the research article (RA), mostly written in English. In the academic context research articles are particularly relevant genres both to students and instructors/researchers, since this is the way in which the researcher is initiated to the academic

community (Árvay–Tankó 2004). In addition, English has become the internationally accepted language of reporting research; therefore, the majority of the investigations have focused on the English conventions of the research article. However, as an increasing number of non-native speakers of English are initiated into the research community, there is a growing need to examine RAs written by non-native speakers to see how they compare with RAs written by native speakers of English. Although there have been some attempts at catering for this need, e.g., Duszak (1994), Golebiowski (1999) and Árvay–Tankó (2004), only the last of these has examined Hungarian writers' texts.

This is all the more unfortunate since in the Hungarian university context, seminar papers, which are essentially theoretical research articles of varying length, are frequently the main bases of evaluation in the humanities. This is also true in the case of students studying English language and literature, who spend a substantial amount of time composing such texts for their tutors. Still, there is little information available on what comprises an acceptable research article. The guidelines that can be accessed are quite short and focus mostly on the appropriate ways of citation.

For example, at one Hungarian university, the guidelines that the Linguistics Department publishes contains a description of what constitutes plagiarism and the expected formal requirements of the papers, e.g., length, papers size, cover sheet format. However, there is no information available on what content elements a successful essay should contain, nor are there sample essays for students to look at. Therefore, English majors usually have two sources to access these necessary pieces of information. For one, they receive training in academic writing during the first year of their studies, where they are taught about the basics of argumentation and the internationally accepted conventions of conducting and writing up research. However, based on student feedback, these conventions are at times not valued by the local teacher community as a number of teachers have their own idiosyncratic requirements for citations and formatting, disregarding international conventions.

The other, though indirect source of information on what constitutes a good piece of writing can be found in the annual/biannual publications of the different English departments, which publish the most successful pieces of student writing in the given year. Since these volumes are edited by the local teacher community, it can be argued that these essays can serve as a good model for students to follow when writing a paper.

At the same time, to date no research has been conducted to see how these student RAs are structured and to what extent they are similar or different from RAs written by native speakers of English who are experienced writers.

Recognizing this need, a study was set up to examine the structure of these papers. As the first step in this project, an analysis of the introduction section of a set of student papers was conducted using Swales' (1990) Create a Research Space (CARS) Model. The purpose of the research was threefold: first, it aimed to see if the papers written in this specific university context observe the typical genre conventions identified by Swales. Second, it aimed to find out if papers written in two different disciplines, literature and linguistics employ different types of introductions. Furthermore, it has been proposed by Árvay–Tankó (2004) that Swales' model was built on the basis of empirical research papers, and that theoretical research papers contain some moves which are not characteristic of empirical papers. The third aim of this research project was to examine the relevance and validity of this claim.

The remainder of the present paper is divided into five parts. First, a summary of the relevant literature will be given, followed by the description of the methodology of the research. The next two sections will introduce the most relevant findings of the project and the conclusions that can be drawn on the basis of these. The last section will summarize the most important findings and indicate pedagogical implications.

2. Literature review

2.1. The analytical tool: Swales' CARS Model

Swales' groundbreaking analytical model came to existence based on research conducted on research article introductions in physics, medicine and social sciences. The first version of this model (Swales 1981) comprised four moves and was based on an analysis of short RA introductions. Due to difficulties in using the analytical model in some contexts, Swales later modified his analytical tool, and the revised CARS model (Swales 1990) comprises 3 Moves: *Establishing a territory, Establishing a niche*, and *Occupying the niche* (see Figure 1 below).

In order to facilitate analysis, Swales gives a detailed description of the structure and function of the specific moves. At the same time, he also includes a list of linguistic examples that characterize each of the steps (for a sample list of these see Appendix A).

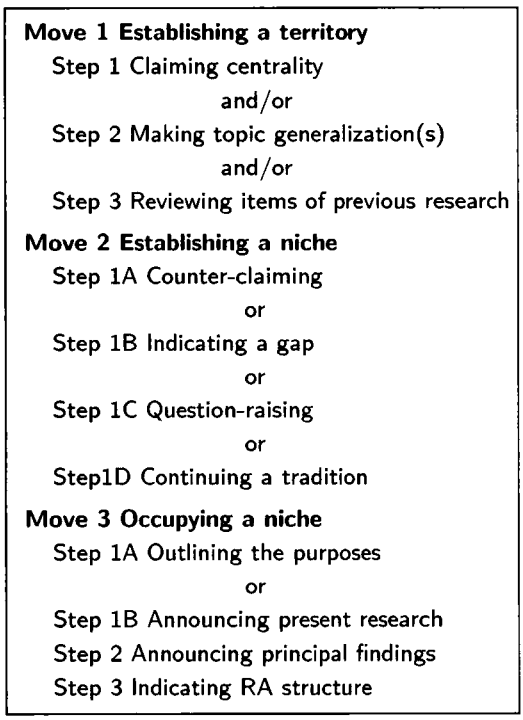


Fig. 1
Swales' (1990) CARS model

Swales argues that the three moves appear in this given order, with under 10% of the introductions beginning with Move 3. On the other hand, he acknowledges that in some cases cyclicity appears: especially in the case of reviewing items of previous research, some authors choose to review individual items separately, followed up with niche-establishment in each of the cases. The only explanation Swales gives for such a strategy is the length of the introductions (longer introductions may tend to use cyclicity more often), or some fields of study tend to value straightforwardness more than others: cyclicity is more typical of social sciences than natural and life sciences and engineering. For establishing territory and establishing a niche (Moves 1 and 2), it is up to the writer to decide which step he or she wants to utilize; however, in Move 3, Step 1 is obligatory, which can be followed by Step 2 and/or 3. When Step 3 is present in the introduction, it is always at the end of it.

2.2. Research conducted with the help of the Swales Model

After presenting his model, Swales proceeds to describe a small-scale analysis of RAs that were published in the field of composition research. The analysis confirms his general claims, although some problems do come to surface. One such problem is the occurrence of Move 3 in an opening position, i.e., when the author begins by outlining the purposes of the study. Although on further investigation of the issue on a different sample Swales found that such an opening strategy is indeed present in some RA introductions, their occurrence was under 20%.

Following Swales' model, substantial amount of research was conducted in different genres, and other parts of the RA (methods, discussion); still, most researchers have concentrated on introductions. Relevant to the present research project is Bunton's (2002) analysis of PhD theses' introductions. Similarly to RAs, dissertations act as an initiation to the research community, and therefore, they are high-stakes "enterprises". Bunton found that nearly all introductions confirmed to Swales' CARS model, although there were some variations as well. Most importantly, the findings indicated that the moves tended to be cyclical, which Bunton explains with the fact that PhD introductions are substantially longer than RA introductions. In this corpus, most introductions began with Move 1, and ended with Move 3. At the same time, Bunton identified much more (in fact, 10 more) types of steps than Swales, most of which occur in Move 3. Most importantly, it was compared whether native and non-native speakers of English differed in their introductions. Bunton found that there was no significant difference between these two groups, and neither was there a major difference between papers written in different disciplines. The present paper analyses the same issues on a different set of papers.

Another important piece of research was conducted by Duszak (1994), who examined language-related RAs written by English and Polish authors. Her findings indicate that though the moves established by Swales do appear in the RAs, their order is not straightforward, since cyclicity appeared in both English and Polish RAs. Another problem was that the Polish texts were often not divided into sections, so it was difficult to determine where a particular "introduction" ended. Duszak also found that there was a tendency for a "strategy of avoidance", i.e., the author's aims were underspecified. In a similar manner, indirectness was typical of the texts, where in Move 3, for example, the writer concluded the in-

troduction with a series of rhetorical questions, without "giving away" his/her position on the issue. In contrast, the English data showed a preference for explicitly stating the purpose of the research. In addition, Polish writers tended to take less responsibility for their findings, and frequently stated what they did not intend to say or accomplish. Duszak's research is important for its cross-cultural focus and its challenging the omnipotence of Swales' CARS model. At the same time, it remains to be seen whether the differences encountered by the Polish writers are carried over to their writing in English as well.

This latter need is recognized by Golebiowski (1999), who analysed scientific article introductions by Polish writers in English and Polish, using the CARS model. Similarly to Duszak (1994), Golebiowski found that Polish articles tended to be unsegmented, which made the identification of introductions troublesome. Only articles written in psychology tended to be segmented. Interestingly, Golebiowski found that Polish papers did not conform to the CARS model at all; only texts written in English provided analysable data. These latter texts showed also some divergence from what Swales identified as a typical sequence. On the whole, Move 1 was replaced by detailed descriptions of background knowledge, and was heavy on literature review. Move 2 was missing in many cases, although Move 3 was present in all of them—though often in a downgraded and indirect way. Golebiowski concludes that the CARS model is not suitable for analysing texts written by Polish writers either in Polish or in English.

The latest contrastive analysis in connection with RA introduction was conducted by Árvay–Tankó (2004). This study is of importance since this has been the only study so far that has analysed Hungarian RA introductions. The research aimed at seeing whether the Swales model applies to theoretical papers as well, and at comparing Hungarian introductions to English ones. Árvay and Tankó identified two new moves in their theoretical RAs: they termed one *Examples*, in which the author illustrates the problem under discussion, usually in the form of italicised lists of words/phrases, or indented blocks of numbered sentences/diagrams. These examples do not illustrate the current state of the art, nor do they review previous research, and their sources are often not indicated. The other Move identified was termed *Analytical details*, in which the writer clarifies the terminology of the RA, much as one does in the Methods section. Furthermore, sometimes a short description of the theoretical framework is also given.

Árvay–Tankó's (2004) contrastive analysis of the two languages has also brought interesting results. The variability of the introductions was much greater in Hungarian RAs, i.e., Hungarian writers had a bigger freedom as to what they can/should include in the introduction. Also, English introductions were much longer and tended to be divided in paragraphs more often than Hungarian ones. There was a tendency among Hungarian writers to raise questions instead of indicating a gap, and principal findings and the description of the RA structure were also often missing. On the whole, it was found that Hungarian introductions differed in their structure from English introductions.

As all of the above indicate, there are several points in connection with the Swales (1990) model that might require further investigation. Most researchers have focused on pieces of writing by native speakers, and it is not clear to what extent the mother tongue of the writers and the disciplines they operate in influence the structure of the introduction. Therefore, the present research set out to examine RA introductions written in a specific university context. The aim of the study was to see to what extent the English-language RA introductions written by non-native students of English conformed to the Swales model and to examine whether there are any rhetorical and structural differences between RAs in two disciplines, linguistics and literature. An additional aim was to investigate if the two new moves proposed by Árvay–Tankó (2004) were present in the current corpus. It was hypothesized that linguistics and literature papers conform to the CARS model, but they employ different moves from one another in their introductions. In addition, it was assumed that the two new moves would be featured by the RAs under investigation.

3. Method

3.1. The corpus

For the sake of the investigation, two sets of texts were collected: a set of ten introductions of linguistics RAs and a set of ten introductions of literature RAs. The motivation for choosing these fields of study was that in the first three years of their studies, English major students need to be able to compose essays overwhelmingly in these areas. As a result, these are the fields of study they need most help in. The texts were chosen from issues published between 1997 and 2003 of an annual journal

of literature and a biannual journal of linguistics, both of which publish the best research articles written by students majoring in English. When selecting the specific text, the researcher made sure that a range of fields within the two areas of study were represented.

In order to have a range of samples, no two introductions were written by the same author. All the authors were non-native speakers of the language (as it was indicated by their names). Furthermore, the beginning and end of the introduction was identified as the authors indicated them: in linguistics the majority of the papers included a subheading 'introduction', although in literature identifying the introductory part was more problematic. In these papers the authors often did not include subheadings at all, but the paper consisted of one continuous stretch of text. Therefore, on the one hand, those papers were chosen which did contain a subheading 'introduction', on the other, those in which the writer divided the texts by empty lines.

3.2. Procedures

After compiling the two corpora, the researcher conducted a first analysis of the texts with the help of the Swales (1990) model, with Árvay and Tankó's (2004) additions. All the texts were analysed into different moves and steps. If a section could not be accounted for by the original Swales model, it was checked whether the two new steps proposed by Árvay and Tankó could account for them. In order to ensure reliability, the same set of texts was re-analysed after a three-week time period, following the same procedures. Altogether 124 steps were identified in the corpus, out of which 5 were coded differently in the two sessions (4%), mostly due to the fact that in some cases a combination of steps were present in the introductions. In 10 cases the researcher found that the particular stretch of text did not conform to any of the steps identified in the two models, these were termed "ambiguous" cases. A sample analysis can be found in Appendix B.

4. Results and discussion

4.1. Comparison of linguistics and literature papers

In terms of the distribution of the moves, there were differences between the two corpora: Move 1 *Establishing a territory* was found much more frequently in the literature RAs than in linguistics ones (35.6% and 18.1%, respectively). Move 2 *Establishing a niche* had approximately the same distribution in the two sets of texts, while there were slight differences in the frequency of Move 3 *Occupying the niche.* For exact numbers see Table 1 below.

Table 1

Distribution of moves in linguistics and literature papers
? = ambiguous cases; QUM3 = quasi-Move 3

Moves	No of steps in LING	% of steps in LING	No of steps in LIT	% of steps in LIT
Move 1	10.5	18.1	23.5	35.6
Move 2	8	13.7	9	13.6
Move 3	29.5	50.8	29.5	44.6
?	6	10.3	4	6
QUM3	4	6.8	0	0

In addition, in both corpora Move 3 was represented most frequently, followed by Move 1 and then Move 2. The total number of steps was 58 in linguistics and 66 in literature, although the total number of words differed to a great extent: the 10 introductions in linguistics added up to 1689 words, while the literature ones totalled 4784. This shows that on average, one Move consisted of 29 words in linguistics, and of 72 words in literature. This latter fact is interesting, since it indicates that longer introductions do not necessarily contain more moves, but it would require further analysis why literature RAs use more words to accomplish essentially the same functions as linguistics papers.

4.1.1. Move 1 *Establishing a territory*

Contrary to Swales' (1990) findings, Move 1 was not found in all introductions: 50% of linguistics and 10% of literature papers did not use it at all. This is also in contrast with Árvay–Tankó's (2004) findings: they found that around 40% of all moves were Move 1s, whereas in the present linguistics corpus this was as low as 18%, although the literature corpus was

closer with 35.4%. As for position, only three linguistics papers started with a Move 1, while seven of the literature papers did so. This may indicate that in linguistics the writers did not feel the need to establish territory, they viewed it as given, while in literature the writers found it more necessary to argue for the importance of their subject matter.

Regarding the steps within Move 1, Step 2A *Making topic generalizations* was used most frequently in the linguistics papers, although this step was used relatively rarely in literature papers (17%). In the literature papers Step 3 *Reviewing items of previous research* was used most. It can be noted, then, that while in linguistics writers deemed it important to make topic generalizations, in literature reviewing items of previous research was estimated to be more important.

Table 2

Distribution of Move 1
S1 = Claiming centrality; S2A = Making topic generalization;
S2B = Examples; S3 = Reviewing items of previous research;
LING = linguistics corpus; LIT = literature corpus

	Nr of steps		Step/Move 1 (%)		Step/corpus (%)	
	LING	LIT	LING (10.5)	LIT (23.5)	LING (58)	LIT (66)
S1	1	4	9.5	17	1.7	6
S2A	5	4	47.6	17	8.6	6
S2B	2	2	19	8.5	3.4	3
S3	2.5	13.5	23.8	57.4	4.3	20.4
TOTAL	10.5	23.5	100	100	18	35.4

4.1.2. Move 2 *Establishing a niche*

Move 2 could not be identified in 60% of linguistics and 50% of literature papers. Where it did appear, it usually took place in the middle of the introduction, in only 2 cases (in 1 linguistics and 1 literature RA) did it appear as the last Move of the introduction. This relative lack of trying to establish a niche suggests that the writers of this corpus tended to avoid challenging other theories or raising questions. As to the specific steps, Step 1B *Indicating a gap* was the most frequent Step in linguistics, while Step 1A *Counter-claiming* was used most often in literature. Though the number of occurrences on the whole is quite low in this corpus, it can be stated that there was a tendency in literature to prefer the direct challenge of other theories to "weaker" options. Linguistics papers, on the other hand, preferred to set minor goals and use softening

more often. At the same time, as it was mentioned above, literature introductions had much more direct reference to previous research, so they had more opportunity to challenge these views than the literature writers who preferred topic generalizations. It was also surprising that contrary to Hungarian traditions, raising questions was not a preferred Step (only one writer used it), and there were no rhetorical questions, either.

Table 3

Distribution of Move 2
S1A = Counter-claiming; S1B = Indicating a gap;
S1C = Question-raising; S1D = Continuing a tradition

	Nr of steps		Step/Move 2 (%)		Step/corpus (%)	
	LING	LIT	LING (8)	LIT (9)	LING (58)	LIT (66)
S1A	2	6	25	66.6	3.4	9
S1B	4	2	50	22.2	6.8	3
S1C	1	1	12.5	11.1	1.7	1.5
S1D	1	0	12.5	0	1.7	0
TOTAL	8	9	100	100	13.6	13.5

4.1.3. Move 3 *Occupying the niche*

The most frequently used Move in both corpora is Move 3: all of the essays have at least one instance, some as many as 5 instances. As Swales (1990) mentions, Move 3 usually occurs as the first or last move: in the present corpus 5 linguistics and 2 literature papers begin with a Move 3, while 9 linguistics and 8 literature introductions end with it. In this respect, this goes against Swales' (1990) findings that under 20% of RA introductions begin with a Move 3. The most frequently used steps were S1A *Outlining purposes* (30.3%) in linguistics and S2 *Announcing principal findings* (44%) in literature.

This latter difference in preferred Moves was what gave a very different feel to the two types of introductions: in literature much space was devoted to describing the findings of the research, i.e., to giving an analysis of the piece of writing under analysis. This description was then detailed in the rest of the paper. At the same time, the writers did not usually state it explicitly that it was their own interpretation they were describing, even if this could be inferred from the context. The linguistics introductions, however, tended to do the opposite: they explicitly described the main features of their research and set very clear goals;

still, it was rare for them to state what the results of their analyses were. This might be due to the fact that there are usually quite a lot of elements in a linguistic analysis and it would be difficult to give an accurate summary of these in one sentence. As for the new S1C *Analytical details*, there were 7 instances of this Step in the linguistics papers, while only 1 instance in literature.

Swales (1990) claims that the obligatory element in Move 3 is Step 1 (*S1A = Outlining purposes; S1B = Announcing present research; S1C = Analytical details).* This was true for all the linguistics papers, all of which contained a Step one, while 4 literature essays did not contain this step, but had a Move 3 Step 2 *Announcing principal findings* instead.

Table 4

Distribution of Move 3

S1A = Outlining purposes; S1B = Announcing present research; S1C = Analytical details; S2 = Announcing principal findings; S3 = Indicating research article structure

	Nr of steps		Step/Move 2 (%)		Step/corpus (%)	
	LING	LIT	LING (29.5)	LIT (29.5)	LING (58)	LIT (66)
S1A	9	8	30.5	27.1	15.5	12.1
S1B	8	5	27.1	16.9	13.7	7.5
S1C	7	1	23.7	3.3	12	1.5
S2	1.5	13	5	44	2.5	19.6
S3	4	2.5	13.5	8.4	6.8	3.7
TOTAL	29.5	29.5	100	100	50.5	44.4

4.1.4. Strategies of avoidance and ambiguous cases

Four cases of avoidance, or "quasi-moves" in Duszak's (1994) terminology, were identified, all of which were found in the linguistics papers. For example, in the sentence *Since this set of data is restricted in some ways, the analysis cannot (and is not meant to) be regarded as definitive or conclusive in any way, only the basic pattern of this harmonic system can be discussed* the author emphasizes what the essay is not doing rather than what it is.

Ambiguous cases were instances when the text did not conform to any of the CARS moves. Contrary to previous expectations, this was only 10% and 6% in the two sets of introductions, which suggests that the model is indeed suitable for analysing these texts. As the numbers

indicate, there was no significant difference between the number of ambiguous cases in the two corpora. The most typical problematic case appeared in literature papers: in several instances the writer gave lengthy descriptions of either the historical era when the piece of literature was written, or the background of certain phenomena mentioned in the text, for example, a lengthy description of what a *bog* means in geographical terms. In other cases the background information "set the scene" for the rest of the introduction, e.g., when the author described who or what inspired him/her to write the paper. The classification of these elements could be the subject of further research.

Table 5

Distribution of quasi-moves

	Nr of steps		Step/QUM (%)		Step/corpus (%)	
	LING	LIT	LING (29.5)	LIT (29.5)	LING (58)	LIT (66)
TOTAL	4	0	NA	NA	6.8	0

Table 6

Distribution of ambiguous cases

	Nr of steps		Step/? (%)		Step/corpus (%)	
	LING	LIT	LING (29.5)	LIT (29.5)	LING (58)	LIT (66)
TOTAL	6	4	NA	NA	10.3	6

4.1.5. Further observations

The order of the Moves was also of interest. Swales (1990) suggests that the usual order is Move 1, Move 2, Move 3. As it was mentioned above, Move 1 indeed tended to come at the beginning of the introduction, while it was usually Move 3 that ended it. However, in 3 linguistics and 4 literature papers Move 1 also appeared in the middle of the essay, usually following a Move 3. Swales also allowed for some cyclicity in the essays, i.e., the repetition of certain moves. This was the case in the present corpus, too: while 7 linguistics and 6 literature papers had a "simple structure" (maximum 3 moves), the rest had "complex" ones, where cyclicity occurred. The fact that 35% of all introductions featured cyclicity indicates that this strategy might be used more often than Swales suggested.

On the whole, establishing territory was more prominent in literature than linguistics papers, as was the need to review items of previous

research. Literature papers tended to challenge other theories more directly than linguistics papers, which used more indirectness. Surprisingly, raising questions or using rhetorical questions was not a preferred strategy. Move 3 was the most frequently used Move in both corpora, but the two fields displayed a preference for different steps: literature papers announced their main findings readily, while linguistics writers tended to outline purposes rather than give away their results.

4.2. Comparison with the results of previous research

The results of the present research confirm those of Árvay–Tankó (2004) at several points. The least frequent Move in both studies is Move 2 *Establishing a niche*, and the new S1C *Analytical details* was utilized in many of the linguistics papers, though not in the literature ones. This tendency may be attributed to the fact that the new Step was proposed based on the analysis of linguistics RAs.

Árvay and Tankó's newly proposed *Examples* Step (S2B), however, was found only in 4 papers, which is a mere 20%. Also, compared to Árvay–Tankó's (*ibid.*) findings, Move 2 Step 1C *Question-raising* was used much more rarely in the present corpus (it constituted only about 1.5% of the corpus), while Step 1A *Counter-claiming* was used much more often. In Árvay and Tankó's corpora Moves 1 and 3 were of more or less equal distribution (Move 1: 39 and 43%, Move 3: 49 and 39%), while in the present case in linguistics Move 1 represented only 18.1% of the total steps.

Our research, on the other hand, is in line with Duszak's (1994) findings in the sense that the order of the moves is not straightforward, but cyclicity appears in many cases. At the same time, it was found that in contrast with Duszak's findings, the field of study influenced the level of directness of the writer, i.e., literature papers tended to present their findings early on, while linguistics introductions only referred to the features of the research. However, it has to be added that indirectness and understatement did occur in the corpus several times.

5. Conclusion

To sum up, the hypothesis that the RA introductions conform to the Swales (1990) model may be confirmed on the basis of the current data-

base, although with some variations. On the whole, even though the majority of the texts could be classified into the Moves/Steps proposed by Swales, the order and distribution of these showed some deviations from Swales' original proposition. Moreover, it was found that the RAs differed based on the field of study in which they were written. Regarding the new Moves suggested by Árvay–Tankó (2004), the use of *Examples* and *Analytical details* could not be justified convincingly. While these moves did appear in the corpus, their frequency was not high. However, the use of *Analytical details* was indeed frequent in linguistics introductions, which confirms the finding that the field of study has an enormous influence on the accepted conventions of the introduction.

Therefore, it is of essential importance that investigations are made to cater for different writers' needs. The analysis of the present corpus, for example, suggests that non-native student writers make use of essentially the same rhetorical elements as those used by expert writers in their field. On the other hand, it is my impression that quite often it is prescriptive intentions that guide researchers in their investigations, in other words, they want to set unified models that are accepted globally. However, it is essential that before this we get to know what is accepted within a genre in different contexts in order to see what different needs there are in different fields of study. To this end, the present project could be followed up by interviewing instructors or editors as to what they consider to be a successful research article. Besides, the success of an essay is only inferred from the fact that it is published in a journal, but different essays or articles might be valued in different contexts. Through this, some of the roots and functions of "strange idiosyncrasies" could be clarified and these could then be incorporated in students' academic writing education.

Appendix A

Examples of textual signs of the rhetorical moves in Swales' (1990) CARS model

Move 1	Step 1: Claiming centrality	*recently, in recent years, great importance, central issue, has been studied by, is a classic problem of, important aspect*
	Step 2: Making topic generalization(s)	*is well known, there is much evidence to support, not completely understood, are often criticized; there are many situations where, it is a common finding that*
	Step 3: Reviewing items of previous research	integral/nonintegral citations
Move 2	Step 1A: Counter-claiming	*however, nevertheless, yet, unfortunately, but*
	Step 1B: Indicating a gap	*suffer, is limited to, time consuming, expensive, not sufficiently accurate*
	Step 1C: Question-raising	direct/indirect questions
	Step 1D: Continuing a tradition	*therefore*
Move 3	Step 1A: Outlining purposes	*this, the present, we, reported, here, now*
	Step 1B: Announcing present research	*the purpose of this investigation is/was to*
	Step 2: Announcing principal findings	
	Step 3: Indicating research article structure	*the paper is divided into five sections, is structured as follows, we have organized the rest of the paper in the following way*

Appendix B

Sample analysis
"Close, But Not Touching", Readings and Misreadings
in John Fowles's *The Collector*

Ever since its publication, John Fowles's *The Collector* (1963) has been a great commercial success –	**M1S1** *Claiming centrality*
"an intriguing study in warped sexuality ... cunningly worked suspense" by "an artist of great imaginative power" –	**M1S3** *Reviewing items of previous research*
as well as the object of intensive critical activity.	**M1S1** *Claiming centrality*
It has been interpreted as a psychological thriller, an allegorical treatment of the struggle between "the Few" and "the Many," a modern version of the Bluebeard legend, a Bildungsroman, an existential journey towards self-discovery, and so on.	**M1S3** *Reviewing items of previous research*
What I want to look at in this study is the issue of interpretation as it is encoded in the novel.	**M3S1A** *Outlining purposes*
In *The Collector* the two protagonists, Frederick Clegg and Miranda Grey enter a reciprocal interpretive game in Clegg's secluded house. It is the nature of this intersubjective reading process that I shall try to explore here. In relation to this, I shall look at the ways the reading process is dramatised within the context of the novel.	**M3S1B** *Announcing present research*
What kinds of reading are approved or rejected by the novel? The most important question proposed by my interpretation is this: is the dichotomy suggested by the novel between apparently good/authentic reading (Miranda) and bad/fake reading (Clegg) still maintained at the end? Finally, is the two characters' interpretation of each other successful – do we have readings or misreadings?	**M2S1C** *Question-raising*

Appendix C

Linguistics

Essay number	Moves									Total Moves
LING1	M1S2A	M1S2B	M1S3	M3S1A	M3S1B	QUM3	M3S1B	M3S3+M3S2		8
LING2	M1S2A	?	M1S2A	M2S1A	M1S3	M2S1B	M3S1A	M3S3+M1S3		8
LING3	?	M3S1A+?	QUM3S1B	M3S1C	M3S2					5
LING4	M3S1C	M2S1B	M1S2A	M3S1C	M1S2A	M2S1B	QUM3S1A	?	M3S1B	9
LING5	M3S1A	M2S1B	M3S1A	?	M2S1C	QUM3S3	M3S1C	M2S1D		8
LING6	M3S1A	M1S2B	M3S1B	M3S1C						4
LING7	?	M3S1A	M3S1B							3
LING8	M3S1C	M3S3								2
LING9	M3S1A	M3S1B	?	M2S1A	M3S1B	M3S3				6
LING10	M1S1	M3S1A	M3S1C	M3S1B	M3S3					5

Essay number	Number of moves	Number of words	Words/Moves	M1	M2	M3	QUM3	?
LING1	8	408	51	3	0	4	1	0
LING2	8	192	24	3.5	2	1.5	0	1
LING3	5	112	22	0	0	3	1	1
LING4	9	133	15	2	2	3	1	1
LING5	8	227	28	0	3	3	1	1
LING6	4	102	25	1	0	3	0	0
LING7	3	154	51	0	0	2	0	1
LING8	2	92	46	0	0	2	0	0
LING9	6	114	19	0	1	4	0	1
LING10	5	155	31	1	0	4	0	0
Total	58	1689	312	10.5	8	29.5	4	6

Literature

Essay number	Moves											Total Moves
LIT1	?	?	M3S2									3
LIT2	M1S2A	M2S1A	M3S1A	M3S1B	M3S3							5
LIT3	M1S2A	M1S3	M1S2B	M1S3	M1S2B	M1S3	M3S1A	M3S1B	M3S1A	M3S1B	M1S3+M3S2	11
LIT4	M3S1A	?	M3S1A	M3S2	M1S3	M3S2						6
LIT5	M3S2	M1S3	M2S1B	M1S3	M3S2	M1S3	M2S1A	M1S3	M3S2			9
LIT6	M1S1	M1S3	M3S1A	M3S1B	M2S1C							5
LIT7	M1S1	M3S1C	M3S2	M3S1A	M3S1B	M3S2						6
LIT8	M1S2A	M1S1	M3S2	M3S1A								4
LIT9	M1S1	M2S1A	M1S2A	M2S1A	M3S2	?	M3S3	M1S3	M3S2			9
LIT10	M1S3	M2S1B	M1S3	M2S1A	M1S3	M2S1A	M3S2	M3S2+M3S3				8

Essay number	Number of moves	Number of words	Words/Moves	M1	M2	M3	QUM3	?
LIT1	3	254	85	0	0	1	0	2
LIT2	5	347	69	1	1	3	0	0
LIT3	11	689	63	6.5	0	4.5	0	0
LIT4	6	253	42	1	0	4	0	1
LIT5	9	873	97	4	2	3	0	0
LIT6	5	218	44	2	1	2	0	0
LIT7	6	338	56	1	0	5	0	0
LIT8	4	394	98	2	0	2	0	0
LIT9	9	1152	128	3	2	3	0	1
LIT10	8	266	33	3	3	2	0	0
Total	66	4784	715	23.5	9	29.5	0	4

References

Árvay, Anett – Gyula Tankó 2004. A contrastive analysis of English and Hungarian theoretical research article introductions. In: International Review of Applied Linguistics in Language Teaching 42 : 71–100.

Bunton, David 2002. Generic moves in Ph.D. thesis introductions. In: John Flowerdew (ed.): Academic discourse, 57–75. Longman, Harlow.

Duszak, Anna 1994. Academic discourse and intellectual styles. In: Journal of Pragmatics 21 : 291–313.

Golebiowski, Zosia 1999. Application of Swales' model in the analysis of research papers by Polish authors. In: International Review of Applied Linguistics in Language Teaching 37 : 231–47.

Swales, John 1981. Aspects of article introductions (Aston ESP research report no. 1). Language Studies Unit, The University of Aston, Birmingham.

Swales, John 1984. Research into the structure of introductions to journal articles and its application to the teaching of academic writing. In: Ray Williams – John Swales – John Kirkman (eds): Common ground: Shared interest in ESP and communication studie, 77–86. Pergamon, Oxford.

Swales, John 1990. Genre analysis: English in academic and research settings. Cambridge University Press, Cambridge.

Acta Linguistica Hungarica, Vol. 53 (2), pp. 117–138 (2006)
DOI: 10.1556/ALing.53.2006.2.2

REMARQUES SUR LES CONSTRUCTIONS CAUSATIVES DU TYPE *FAIRE* + INFINITIF DANS LES LANGUES ROMANES ET QUELQUES AUTRES

MAREK GAWEŁKO

ul. Włoska 19/20
30-638 Kraków
Pologne

Résumé : L'objectif de l'article est double : 1° établissement de la vitalité des constructions causatives dans huit langues dont deux (l'anglais et le français) sont hautement analytiques, une (le polonais) hautement synthétique et cinq qui montrent une tendance analytique modérée, et 2° explication typologique des différences observées.

La règle générale dit que la vitalité des constructions causatives est en rapport avec le degré d'analytisme de la langue où elles sont utilisées : elles sont les plus fortes en français et inexistantes en polonais. L'anglais fait exception à la règle : les constructions causatives y sont relativement rares, ce qui est en rapport avec la faiblesse relative de l'infinitif, concurrencé par les formes en *-ing*.

Mots clés : causatif, vitalité, explication, synthétique, analytique

1. Remarques générales

Les formes linguistiques porteuses d'une valeur causative ont été établies pour différentes langues du monde, y compris les langues romanes (cf. par ex. Shibatani 1976b; Givón–Young 2002; Shibatani 2002b; Iliescu 1995; Moeschler 2003, etc.). Les formes causatives les plus importantes sont : le causatif lexical (ex. *tuer*), le causatif morphologique (ex. *fertiliser*), le causatif syntaxique, par ex. du type *faire* + inf(initif). On indique aussi d'autres formes : participe, gérondif, constructions telles que préposition + inf, préposition + substantif, etc. (cf. Mori 1980, 30, note 13). Shibatani et Pardeshi (2002, 109) signalent des procédés tels que réduplication et

1216–8076/$ 20.00

apophonie. Il y a sans doute un rapport entre l'importance accordée aux types lexical, morphologique et syntaxique et le caractère typologique des langues où ils apparaissent.

Le causatif lexical est le moyen le plus fréquent d'exprimer la causalité. Cependant, il pose problème pour le linguiste vu que seul un critère sémantique permet de le distinguer du verbe non causatif. De plus, le caractère causatif du verbe devrait être établi dans le contexte où il est utilisé. Ainsi *open* n'est causatif que dans la deuxième phrase: *The door opened. John opened the door* (cf. Shibatani 1976a, 3). Il en va de même du verbe français *aimer*, qui n'est pas causatif. Cependant il le devient dans certains contextes, par ex. *il l'aime à en être fou* (Sandfeld 1965, 249). La délimitation des verbes causatifs peut être facilitée quelque peu par la description des « situations causatives » (cf. Shibatani 1976a, 1).

Le causatif morphologique se caractérise par la présence de morphèmes causatifs. Par ex. pour l'espagnol, Aranda Ortiz (1990, 159 sq) établit les suffixes *-ificar (dulcificar), -izar (fertilizar), -ar (fecundar), -ecer (acontecer)*, etc. Le causatif morphologique est relativement fréquent dans les langues synthétiques. Ainsi, en hindi il montre une grande vitalité, comme en témoigne la constatation de Pobożniak (1979): « Das Causativum ist in der Hindi-Sprache eine reichlich entwickelte und produktive Kategorie, so daß man fast zu jedem Zeitwort ihre Form bilden kann. » Une telle vitalité n'est pas toujours le cas. Les causatifs morphologiques slaves, par exemple, sont différenciés, mais leur fréquence semble être beaucoup plus basse. Ainsi, Gołąb parle de trois types de causatifs slaves (1968, 93) qui se caractérisent par 1° un suffixe et une apophonie, 2° un suffixe, 3° une déréflexivisation. Cependant, cette fréquence est relativement considérable, comme en fait foi l'existence de certains exemples considérés comme impossibles par des spécialistes. Ainsi Shibatani (2002b, 3) dit: « We are likely not to find a language in which causatives corresponding to verbs such as 'swim', 'sing' [...] are lexicalized ». Pourtant de tels causatifs apparaissent dans les langues slaves, cf. pour le slovaque l'exemple de Heinisch (1977), possible aussi en polonais: *Tento spevák vie rozespevat' celú salú* 'Il sait faire chanter toute la salle'.

Les types morphologique et syntaxique servent à suppléer à la faible extension du lexique causal. De plus, le dernier type, permet « d'introduire une relation causale forte ou faible par la simple introduction d'un agent en position de sujet de l'opérateur *faire* ou *laisser* » (Moeschler 2003, 14). La construction en *laisser*, ou permissif, est souvent considérée comme un causatif à ceci près que sa force causative est faible (cf. p. ex. Simone –Cerbasi 2001, 446).

2. La tâche

La majorité des travaux sur les constructions causatives concerne la syntaxe (ex. le passage d'un verbe intransitif au transitif), plus rarement la sémantique (ex. la causalité directe vs indirecte; cf. aussi Reboul 2003), rarement l'aspect communicatif (cf. Iliescu 1995, 3). On a étudié aussi l'aspect historique des constructions causatives romanes (cf. Simone–Cerbasi 2001).

Dans notre article, nous nous limitons à la construction de base: *faire* + inf et incluons son synonyme fréquent et partiellement grammaticalisé dans les langues ibéro-romanes *mandar (envoyer)* + inf.

Nous nous assignons deux objectifs principaux. En premier lieu, nous essayons de déterminer la vitalité des constructions causatives en *faire* utilisées dans cinq langues romanes (l'espagnol, le français, l'italien, le portugais, le roumain), deux langues germaniques (l'allemand, l'anglais) et une langue slave (le polonais). Nous avons établi aussi la vitalité des équivalents, causatifs (lexicaux et morphologiques) et non causatifs, de ces constructions. Notre description se pose pour but 1° de différencier les cinq langues romanes et 2° de les opposer aux trois langues non romanes.

En deuxième lieu, nous essayons de déterminer à quel point la vitalité des constructions mentionnées s'explique par le caractère typologique des langues où elles sont utilisées.

Le premier de ces objectifs est moins important pour deux raisons: 1° l'analyse des causatifs romans a fait l'objet de plusieurs études (p. ex. Iliescu 1995; Simone–Cerbasi 2001) et 2° l'explication est plus intéressante que la description.

3. Le corpus

Notre corpus embrasse cinq langues romanes: l'espagnol, le français, l'italien, le portugais et le roumain, deux langues germaniques: l'allemand et l'anglais, une langue slave: le polonais.

Le corpus I comporte presque 400 000 mots graphiques pour chaque langue. Il se compose de 6 ouvrages dans leur intégrité (4 ouvrages français: CH, E, P, V, 1 ouvrage anglais: A et 1 ouvrage latin: IC), des trois premiers chapitres d'un ouvrage portugais (ML) et de fragments de deux ouvrages polonais (CD et Q). Pour chaque ouvrage, nous disposons de l'original et de traductions en sept langues. Une exception: le texte CD n'a pas de version portugaise, ce qui nous a forcé à prendre en compte un

fragment de la version portugaise d'un autre ouvrage (BA) de longueur équivalente à celle du texte manquant.

Trois de ces ouvrages sont rédigés dans une langue proche de la langue parlée : A, ML et CD. La langue de IC est quelque peu archaïsante.

Comme il n'est pas facile de manier cette masse d'informations, nous avons constitué un autre corpus, défini comme corpus II, utilisé surtout au tableau 1. Il se compose, pour chaque langue, de fragments des neuf ouvrages indiqués plus haut. De chaque ouvrage, nous avons retenu le fragment initial et le fragment final, chacun comportant, dans la version française, 2500 mots graphiques, à ceci près que nous ne coupons pas les phrases. Le corpus français comprend ainsi 45 057 mots graphiques, au lieu de 45 000 si les phrases avaient été coupées. Le corpus français une fois établi, nous avons constitué aussi le corpus des autres langues qui est à chaque fois composé des fragments équivalents. Ainsi, le corpus II comporte 45 000 mots environ, quantité qui caractérise un roman moyen (cf. CH : presque 31 000 mots, V : 76 500 mots environ). Cependant, il ne permet pas de traiter certains problèmes, par ex. celui du causataire introduit par *à* ou *par* (cf. Marstrander 2000, 205).

4. Occurrences des constructions causatives et de leurs équivalents

Les verbes causatifs sont nombreux. Ainsi pour le polonais, Termińska (1983, 18) a accumulé un corpus de 500 verbes causatifs (= causatifs lexicaux) environ. Les occurrences des constructions causatives du type *je fais le chien aboyer* sont beaucoup plus rares. Pour rendre compte de la vitalité de ces constructions, nous nous fondons sur les occurrences qu'elles présentent dans des textes parallèles, tandis que la richesse des fonctions syntaxiques qu'elles assument, qui différencient peu les langues romanes, sera mentionnées plus bas (§ **6.4**).

4.1. Tableau des occurrences

Le tableau 1 rend compte des constructions causatives du type *faire* + inf (roum. *face* + subjonctif) et de leurs équivalents : 1° causatif lexical et causatif morphologique et 2° équivalent non causatif. Nous retenons aussi l'équivalent ibéro-roman approximatif partiellement grammaticalisé du type *mandar* + inf ; la valeur causative s'y confond en général avec la

valeur lexicale de ce verbe (cf. Hernanz Carbó 2000, 2263). Par contre, nous n'avons pas retenu les verbes tels que prt. *forçar*, ang. *bring, order*, all. *befehlen*, roum. *porunci*, pol. *kazać*, fr. *ordonner*, etc., p. ex. *il m'ordonne de partir* au lieu de *il me fait partir*. Même en tant qu'équivalents du verbe *faire* dans sa fonction causative, ils sont tout au plus partiellement grammaticalisés, leur sémantique reste concrète.

Le verbe esp. prt. *mandar* est plus grammaticalisé que les verbes cités ci-dessus. On peut avancer un argument grammatical et un argument sémantique. Le premier c'est avant tout le comportement du pronom : il apparaît à l'accusatif avant *mandar*, tandis que devant des verbes tels que *ordenar* il apparaît au datif (pour plus de détails, cf. Hernanz Carbó *op.cit.*, 2263–5). Un phénomène analogue se manifeste en principe aussi en portugais, p. ex.

(1) (a) fr. Le baron venait de *le faire remettre* à neuf (V, 35)
(b) prt. barão acabava de o *mandar arranjar* de novo

(2) (a) prt. Só *o mandarão entrar quando* (E, 137)
(b) prt. onde *me mandou sentar* (E, 95)

Le deuxième, c'est la sémantique généralisée de *mandar* qui apparaît par ex. dans

(3) (a) fr. On voulut sonner le domestique pour *faire avancer* la voiture. (V, 117)
(b) esp. Quisieran llamar el criado para que *hiciese adelantar* el carruaje.
(c) prt. Quiseram tocar, a chamar o criado, para que este *mandasse avançar* a carruagem.
(d) it. Si volle richiamare il servitore per *far avvicinare* la carrozza.

(4) (a) prt. Alguns passos adiante, *mandam-me parar* e perguntam-me se aquela mulher é minha Mãe. (ML, 81)
(b) fr. Quelques pas plus loin, ils me *firent arrêter*.
(c) esp. Algunos pasos más adelante, me *mandaron parar* y me preguntaron.

Il y a un certain nombre de différences entre la construction *mandar* + inf en espagnol et en portugais.

Une différence qui saute aux yeux c'est la fréquence d'emploi. La construction du type *mandar* + inf apparaît 1 fois dans les versions espagnoles de A, de P, de IC et de V, 0 dans CH et dans E. Il en va tout autrement des versions portugaises : 19 occurrences dans V, 17 dans P,

10 dans A, 9 dans E, 2 dans CH, 1 dans IC. Dans ML il s'agit probablement de l'influence de l'original sur la traduction: 47 occurrences dans l'original portugais et 46 en espagnol.

Le subjonctif après *mandar* peut apparaître dans les deux langues, il est plus fréquent en espagnol, par ex.

(5) (a) prt. Mandei-a, então, dar-lhe um abraço e um beijo. (ML, 22)
(b) esp. Le mandé *que le diera* un abrazo y un beso.

(6) (a) fr. Malgré le froid, la baronne fit apporter un siège afin de le regarder travailler. (V, 113)
(b) prt. Apesar do frio, a baronesa mandou *que lhe trouxessem* uma cadeira.
(c) it. Nonostante il freddo, la baronessa fece portare una sedia.

Ainsi, il y a un certain rapport entre la quantité des constructions causatives et celle des subjonctifs (cf. aussi Iliescu 1995, 11).

Pour le roumain, on signale, à côté du tour *face* + subjonctif, aussi la construction *pune* + subjonctif (cf. *idem.*). Ces tours sont en grande partie grammaticalisés. Cependant le degré de grammaticalisation est moins élevé en roumain que dans les autres langues romanes. Pour certaines langues on se demande s'il n'est pas légitime de parler d'un seul verbe causatif, par ex. Comrie constate (1976, 296): «the analytic causative in certain languages, such as the French causative with *faire* [...] functions in many respects like a fused verbal construction.» Or, une telle constatation n'est pas motivée pour le roumain où le verbe *face* est suivi d'une conjonction qui lui assure une certaine autonomie. Un objet sépare facilement le verbe *face* et le subjonctif, comme dans

(7) (a) fr. Arrivé là, on admettra sans peine que rien ne pouvait faire espérer à nos concitoyens les incidents qui se produisirent au printemps de cette année-là. (P, 7)
(b) esp. Siendo así las cosas, se admitirá fácilmente que no hubiese nada que hiciera esperar a nuestros conciudadanos los acontecimientos que se produjeron a principios de aquel año.
(c) it. Qui giunti, si ammetterà senza fatica che nulla poteva far presagire ai nostri concittadini gli incidenti che si verificarono nella primavera di quell'anno.
(d) prt. Uma vez lá, admitir-se-á sem custo que nada podia fazer esperar aos nossos concidadãos os incidentes que se produziram na Primavera desse ano.
(e) roum. Ajunşi aici, vă fi lesne să admitem că nimic nu i putea face pe *concetăţenii noştri* să prevadă incidentele care au avut loc în primăvara anului acela.

Le tour roum. *pune* + inf est présenté par Iliescu (1993). Il ne ressort pas de ce travail que la construction du type *face* + subjonctif jouit d'une richesse comparable à celle des constructions ibéro-romanes en *hacer/ fazer*, tandis que le tour roum. en *pune* apparaît sporadiquement. Ainsi dans notre corpus II, il apparaît 3 fois à peine (la construction en *face* 25 fois). Dans V on a 68 constructions en *face* et 6 en *pune*. Dans E, on trouve 6 constructions en *pune*, dans A—5, dans CH—2.

Les langues germaniques offrent quelques constructions causatives. Duffley (1992, 57) parle des constructions anglaises en *make, cause* et *have*. Or, leurs occurrences dans notre corpus sont peu nombreuses: la construction en *cause* apparaît 12 fois dans tout notre corpus I (7 dans IC, p. ex. *the love of Jesus causes man to regard himself very humbly*), celle en *have* n'apparaît jamais comme équivalent de fr. *faire* causatif dans V. Il en va de même de all. *heißen* qui forme trois constructions causatives dans notre corpus I (ex. *führte man mich mit dieser Dame in einen Saal und hiess uns ein wenig warten*—ML).

Tableau 1

Occurrences des constructions causatives introduites par fr. *faire, envoyer*; esp. *hacer, mandar*; it. *fare*; prt. *fazer, mandar*; roum. *face*; ang. *make*; all. *lassen* et de leurs équivalents dans le corpus II

Langue / Forme	FR	ESP	IT	PRT	ROUM	ANG	ALL	POL
Construction *faire* + inf	48	28	39	26	*28	12	30	0
Causatif synthétique	13	18	13	13	9	24	8	27
Construction *envoyer* +inf	2	5	1	8	2	3	2	9
Total formes causatives	63	51	53	47	39	39	40	36
Équivalent non causatif	12	24	22	28	36	36	35	39
Total	75	75	75	75	75	75	75	75

*—le tour roum. *face/pune* + subjonctif; causatif synthétique = causatif lexical et, très rarement, causatif morphologique

Comme au tableau 1 nous n'avons pas retenu tous les verbes causatifs, mais seulement ceux qui ont pour équivalent au moins dans une autre langue une construction causative, le tableau ne permet pas de déterminer le rôle qui revient à trois types de causatif: lexical, morphologique et

syntaxique dans l'expression de la causalité. Cependant, il jette quelque lumière sur ce problème vu qu'on a tenu compte aussi des deux types d'équivalents: causatif et non causatif. Nos conclusions les plus importantes sont les suivantes. La tendance générale à envisager une action comme causative est la plus forte en français. Elle est un peu plus faible dans les autres langues romanes, le roumain excepté. Elle est la plus faible en roumain et dans les langues non romanes.

Les données relatives à la construction allemande en *lassen* sont peu sûres, le même auxiliaire servant aussi à former le permissif.

L'emploi de la construction causative est facilitée par la mise en relief de la valeur causative, comme dans

(8) (a) fr. S'il s'en est servi, c'est seulement pour *comprendre ou faire comprendre* ses concitoyens et pour donner une forme, aussi précise que possible, à ce que, la plupart du temps, ils ressentaient confusément. (P, 249)

(b) esp. Si se sirvió de ella fue solamente por *comprender o hacer comprender* a sus conciudadanos.

(c) it. S'egli se n'è servito, è soltanto per *capire o far capire* i suoi concittadini.

(d) prt. Se se serviu delas, foi apenas para *compreender ou fazer compreender* os seus concidadãos.

(e) roum. Dacă s-a folosit totuşi de ele, a fost numai ca *să-i înţeleagă sau să facă să fie înţeleşi* concetăţenii.

(f) pol. Tylko po to, żeby rozumieć lub wytłumaczyć swych współobywateli.

Seul le polonais oppose les deux sens, causatif et non causatif, sans dépasser le niveau lexical.

En ce qui concerne la construction du type *faire* + inf, seul le français se sépare nettement des autres langues, relativement peu de l'italien. Les langues qui restent se divisent en trois catégories: 1° l'espagnol, le portugais, le roumain et l'allemand, 2° l'anglais et 3° le polonais. Les langues appartenant à la première catégorie sont plus pauvres que les langues « du centre » : le français et l'italien. La prise en compte des rares exemples de esp. prt. *mandar* + inf ne change pas la conclusion générale. La faiblesse relative de la construction causative anglaise devrait étonner vu le caractère analytique de cette langue. La construction du type « accusativus cum infinitivo », dont la construction causative est un cas particulier, n'a pas survécu en polonais et les tentatives de la réintroduire ont échoué (cf. Urbańczyk–Kucała 1999, 11).

4.2. Deux autres enquêtes

En vue de parer aux faiblesses du corpus II, nous procédons encore à deux enquêtes destinées à confirmer la vitalité des constructions en *faire*. La première (tab. 2) est fondée sur deux ouvrages entiers: A et E. Ici, nous n'indiquons pas la construction roumaine du type *face* + subjonctif.

Tableau 2

Occurrences des constructions causatives romanes en *faire* (aussi *envoyer*, esp. prt. *mandar*) + inf sur la base de deux ouvrages

Forme	Langue Œuvre	FR	ESP	IT	PRT	ROUM
Construc.	A	51	16	35	20	0
causative	E	32	30	38	12	0
Total		83	46	73	32	0

Les données réunies au tableau confirment celles du tableau 1. Les constructions causatives françaises et italiennes sont plus riches que les constructions causatives espagnoles, portugaises ou roumaines. Les formes roumaines *face* + inf ne sont pas représentées dans les deux tableaux, mais Iliescu (1995, 11) constate que l'infinitif, au lieu du subjonctif, peut apparaître exceptionnellement.

La deuxième enquête porte sur le roman *Une vie* de Maupassant: 116 constructions causatives françaises en *faire*, 76 italiennes, 68 constructions roumaines du type *face* + subjonctif et 2 du type *pune* + subjonctif, 119 constructions allemandes en *lassen*, dont 70 environ sont des causatifs propres (les autres, des permissifs). Elle a confirmé aussi la faiblesse de la construction anglaise *make* + inf: dans les premières 1000 phrases de V, on trouve 33 constructions françaises en *faire* et seulement 12 anglaises en *make*.

L'enquête est motivée entre autres par l'opinion avancée par Iliescu (1995, 12): « Pour le roumain, l'emploi de *a face* comme auxiliaire factitif est souvent exclu ». Or, notre enquête, qui ne concerne pas les formes acceptables mais seulement utilisées dans un texte, apporte une certaine contribution à ce problème: numériquement, les constructions roumaines ne s'avèrent pas plus pauvres que les constructions ibéro-romanes et dans V, elles sont aussi riches que les constructions italiennes. Iliescu aurait dû préciser les catégories de constructions françaises ou italiennes dont les équivalents roumains fidèles sont impossibles. Notre corpus permet de

signaler l'absence du tour roumain équivalent au tour fr. *se faire* + inf, qui ne semble pas être causatif, par ex.

(9) (a) fr. Le froid du matin commençait à *se faire sentir* dans la pièce. (P, 235)
(b) esp. empezaba a *hacerse sentir* en la habitación.
(c) it. Il freddo della mattina cominciava a *farsi sentire* nella stanza.
(d) prt. O frio da manhã começava a *fazer sentir-se* no quarto.
(e) roum. Frigul dimineții începea *să se facă simțit* în odaie.

Un autre problème c'est la vitalité de la construction causative italienne du type *fare* + inf. Différents chercheurs constatent une similitude poussée entre le français et l'italien. Ainsi Iliescu (1995, 12) écrit : « le français et l'italien se comportent d'une façon étonnante presque de la même façon, ce qui est confirmé par la totalité des exemples trouvés dans la traduction du livre de Simenon ».

À leur tour, Simone et Cerbasi (2001, 441) disent : « Italian and French are strongly causative-oriented languages, i.e., they tend to see facts as causatively structured to an extent that is probably unique among the Romance languages. »

En dépit de ces opinions, nos tableaux montrent qu'il y a une différence numérique entre les constructions causatives dans les deux langues. Aussi bien notre tableau 1, basé sur 9 ouvrages dont 4 seulement sont français (48 ex. fr. contre 39 ex. it.), que les données offertes par un ouvrage anglais (A : 51 ex. fr. contre 35 ex. it.) et par un ouvrage français (V : 116 ex. fr. contre 76 ex. it.) mettent en évidence une différence nette entre les deux langues. Les données différentes offertes par un seul ouvrage (E : 38 constructions causatives italiennes contre 32 françaises) ne changent pas la conclusion générale. Cette conclusion est corroborée par des facteurs autres que numériques (cf. ci-dessous, § **5**).

5. Causalité directe et indirecte

Une définition très simple de Shibatani et Pardeshi (2002, 88) nous suffit : « Therefore it is a good first approximation to define direct causative and as a situation involving an agentive causer and a patientive causee and indirect causation as one involving two agentive participants, one an agentive causer and the other an agentive causee. »

En principe, le verbe simple désigne une causation directe, la construction en *faire*—une causation indirecte. Cependant, la causation lexicale est ambiguë vu que le verbe simple peut dénoter assez souvent aussi bien une causation directe qu'indirecte: *J'ai réparé ma voiture* peut signifier soit 'je l'ai réparée moi-même' soit 'j'ai fait réparer ma voiture par un mécanicien'. La construction causative en *faire* est ambiguë elle aussi vu qu'elle peut dénoter aussi bien une causation indirecte que directe (ex. *il fit remarquer*—1° *il dit*, 2° *il fit qu'un autre remarqua*). Une autre ambiguïté est signalée par Robustelli (1995, 203). Une phrase telle que *Faccio portare il fratello* peut avoir deux interprétations: 1° *Faccio che il fratello porti* et 2° *Faccio che il fratello sia portato.*

Dans un certain nombre d'exemples, seul le français exprime une causalité indirecte, c'est-à-dire qu'il emploie plus de constructions du type *faire* + inf que les autres langues prises en compte, y compris l'italien. Dans l'exemple qui suit, aussi bien l'original polonais que les traductions emploient un verbe simple même si la causalité relative à l'action d'enlever est indirecte. Seul le traducteur français introduit la construction en *faire*, plus appropriée dans ce cas:

(10) (a) pol. Zamiast pokłonić się o nią Aulusowi i Pomponii, *odebrał* dziecko podstępem rodzicom. (Q, 88)

(b) fr. Au lieu de se présenter devant Aulus et Pomponia pour la leur demander, il *l'avait fait enlever* par subterfuge.

(c) esp. *les había arrebatado* la hija valiéndose de astucias

(d) it. *aveva strappato* con un inganno la figlia ai genitori

(e) prt. *arrebatara-a*, de surpresa, aos seus pais adoptivos

(f) roum. o *răpise* prin surprindere, părinților ei

(g) all. *hatte* er sie *entführt* durch List

Dans cet exemple, la connaissance du contexte précédent permet de constater que Vinicius n'a pas enlevé lui-même la femme en question, par conséquent l'emploi de la construction causative en français est plus régulier que l'emploi du verbe simple dans les autres langues qui procèdent à l'« effacement en surface ».

Dans l'exemple suivant, où il s'agit d'une causalité directe, l'original portugais présente un causatif lexical, qui est retenu dans les traductions sauf en français:

(11) (a) prt. E quem melhor que este Imaculado Coração nos poderia *descobrir* os segredos da Divina Misericórdia? (ML, 16)

(b) fr. Et qui d'autre que le Coeur Immaculé pourrait *faire découvrir* les secrets de la Miséricorde Divine ?
(c) esp. Y, ¿quién mejor que este Inmaculado Corazón nos podría *descubrir* los secretos de la divina Misericordia ?
(d) it. E chi meglio di questo Cuore Immacolato ci potrebbe *scoprire* i segreti della Misericordia Divina ?
(e) roum. Şi cine alta decât Neprihănita Inimă ar putea *să dezvăluie* tainele îndurării divine ?
(f) ang. could *have revealed* to us the secrets of Divine Mercy ?
(g) pol. *odkryć* tajemnice Boskiego Miłosierdzia ?
(h) all. die Geheimnisse der göttlichen Barmherzigkeit *aufdecken* ?

Dans presque toutes ces versions, il s'agit d'une seule action (de découvrir certains secrets) accomplie par un seul agent et non pas d'inciter un autre agent à le faire. Seul le français emploie une construction du type *faire* + inf et envisage l'action comme une incitation à découvrir certains secrets. Dans les autres langues cela se produit plus rarement. Ici il convient de signaler des tours tels que *faire remarquer, faire observer* qui, en dépit de leur caractère analytique, se rapportent souvent à une causalité directe.

Ces formes analytiques sont possibles non seulement en français, mais dans cette langue elles sont les plus fréquentes, par ex.

(12) (a) fr. Il aurait pu recommencer, comme le lui *fit remarquer* Rieux. (P, 71)
(b) esp. como le *decía* Rieux
(c) it. come gli *osservò* Rieux
(d) prt. com lhe *fez notar* Rieux
(e) roum. cum l-*a făcut să observe* Rieux
(f) ang. as Rieux *pointed out*
(g) all. wie Rieux es ihm *riet*
(h) pol. ak *zauważyl* Rieux

Les langues non romanes emploient des équivalents synthétiques ; voir encore

(13) (a) fr. Mais d'autres sont morts, *fit remarquer* Rieux. (P, 44)
(b) esp. *hizo observar* Rieux
(c) it. *fece notare* Rieux
(d) prt. *fez notar* Rieux
(e) roum. *atrage atenţia* Rieux

(f) ang. Rieux *observed*

(g) all. *bemerkte* Rieux

(h) pol. *zauważył* Rieux

L'italien est un peu plus pauvre que le français. Une différence se manifeste quand l'original n'est pas roman. Dans A on trouve 15 exemples français (10 du type *faire observer* et 5 du type *faire remarquer*) et 0 italiens, dans Q—2 exemples français et 0 italiens, mais dans CD 0 pour les deux langues.

Ainsi, la réalisation de l'opposition causalité directe/causalité indirecte apporte un argument de plus en faveur de la thèse qu'on ne devrait pas traiter sur un pied d'égalité le français et l'italien (cf. aussi ci-dessous, § **6.1**) comme on le fait parfois (cf. Iliescu 1995, 12). Notre corpus conduit à différencier les deux langues.

6. Remarques typologiques

Les langues romanes sont liées non seulement par l'origine commune, mais aussi par un certain complexe de traits typologiques. Rien d'étonnant à ce que certains parlent d'un type roman (par ex. Coşeriu 1988). Nous avons montré à plusieurs reprises que les langues romanes se distinguent des langues non romanes prises en compte ici par le degré de synthétisme/analytisme de cette façon que le polonais et, dans une mesure plus limitée, l'allemand sont plus synthétiques, l'anglais—plus analytique que les langues romanes (cf. Gawełko 2001, 40). Il n'est pas sans intérêt de déterminer si la construction analytique en *faire* (all. *lassen*, ang. *make*) est effectivement plus riche en anglais et moins riche en allemand et en polonais que dans les langues romanes.

On peut généraliser cette tâche et proposer de déterminer le rapport entre la vitalité de la construction en *faire* et la tendance analytique. Une telle tâche est légitime vu que des spécialistes de typologie universelle présentent les formes causatives dans le cadre d'une échelle qui va de la forme synthétique vers la forme analytique, comme dans Givón–Young (2002, 52): lexical causative (*She broke the window*), morphological causative (*She enlarged the house*), co-lexicalization (*She let-go of his hand*), non-finite complementation (*She made him leave*), subjunctive complementation (*She asked that he leave*). Ils constatent à propos de cette série de formes: « The causative constructions at the top of

the complementation scale display the highest level of clause integration, code the strongest and the most direct causation, and are the most likely to involve a non-human patient as the causee. Toward the bottom of the scale, constructions display lower levels of clause union, code weaker and/or less-direct causation, are more likely to have a human agent as the causee» (*op.cit.*, 51).

Notre corpus II ne peut que confirmer ce schéma, notamment en ce qui concerne l'agent 1° d'un causatif lexical et 2° d'un infinitif introduit par *faire*. Or, le causatif lexical comporte un agent animé ou inanimé dans des proportions rapprochées, tandis que dans la construction en *faire* l'agent animé de l'infinitif prédomine largement : 30 agents animés contre 18 inanimés dans le texte français, 8 animés contre 4 inanimés dans le texte anglais.

Dans les langues du monde une distinction semble se dessiner : le causatif analytique apparaît de préférence dans les langues analytiques tandis que les langues synthétiques ont recours à des affixes (cf. Shibatani 1976b, 2–3). Ce principe trouve une confirmation en latin où, à en croire D. Strong (cf. Simone–Cerbasi 2001, 450), on observe un passage de causatifs morphologiques aux causatifs syntaxiques.

La question se pose de savoir si, dans les langues prises en compte, la distribution des formes causatives les plus importantes est conforme au caractère typologique de ces langues.

6.1. Le français

La forte tendance analytique du français semble expliquer la grande vitalité des constructions causatives en *faire*, observée aux §§ 4 et 5. Cette tendance explique donc les occurrences de la construction française *faire* + inf et les cas particuliers : 1° l'emploi régulier de cette construction dans le cas d'une causalité indirecte, tandis que les autres langues tendent à employer dans ce cas un causatif synthétique (lexical) et 2° l'emploi de la construction *faire* + inf dans le cas d'une causalité directe, phénomène plus rare dans les autres langues.

6.2. Les autres langues romanes

À lire ce qui précède, un doute surgit : l'italien comporte une construction causative analytique qui se rapproche de celle en français et pourtant

sa tendance analytique est beaucoup plus faible que celle du français. Or, la marche vers l'analytisme, qui caractérise bien les langues indo-européennes, n'est pas un processus homogène. L'infinitif italien évolue relativement vite. Telle sa tendance nominale: au cours de son histoire l'infinitif italien gagne en caractère nominal. Les informations apportées par Vanvolsem (1983, 168 tab.) confirment cette remarque: l'infinitif est sensiblement plus nominal chez Manzoni (XIX^{e} s.) que chez Boccaccio (XIV^{e} s.).

Cependant, cet argument a une valeur très limitée. La force de la tendance analytique est très rapprochée dans les langues espagnole, italienne et portugaise, tandis que la vitalité de la construction en *faire* différencie ces langues d'une façon notable: l'italien se sépare nettement des langues ibéro-romanes en se rapprochant du français.

L'explication de la forme roumaine est relativement simple. La construction causative la plus fréquente *face* + subjonctif est en rapport avec la faiblesse de son infinitif, laquelle est un trait balkanique. Ainsi, c'est avant tout la géographie linguistique qui explique la tendance à remplacer l'infinitif par le subjonctif.

6.3. L'allemand

L'explication de la fréquence relativement considérable de la construction allemande est plus difficile. Une explication partielle est que la tendance analytique de l'allemand est plus forte qu'elle ne paraît de prime abord. La déclinaison des substantifs, qui donne l'impression d'une langue hautement synthétique, y est faible. L'opposition entre le nominatif et l'accusatif ne persiste qu'au masculin singulier. On trouve des formes caractéristiques des langues analytiques, par ex. l'article, le pronom sujet obligatoire, le présentatif du type *es gibt*. On note de plus la faiblesse de l'aspect verbal, la force du passif, etc. (cf. Gawełko 2001, 29). Ainsi, la fréquence considérable du causatif analytique en *lassen* ne devrait pas nous étonner excessivement.

6.4. L'anglais

L'irrégularité la plus importante c'est la faiblesse relative de la construction anglaise du type *make* + inf. Avant d'en proposer une explication, il faut prendre position par rapport à certaines opinions. D'après cer-

tains chercheurs, l'infinitif anglais en général aussi bien que sa proposition infinitive sont forts. Ainsi Rémi-Giraud est d'avis que, à l'opposé du français et de l'allemand, « l'anglais et le grec connaissent une grande extension de la proposition infinitive » (1988, 59). L'argument est apporté par la richesse du « système d'opposition aspectuelle ». L'anglais présente un « système à quatre formes » (*to make, to have made, to be making, to have been making*). Par contre, ne disposant que de deux formes (fr. *dormir* et *avoir dormi*, all. *schlafen* et *geschlafen haben*), l'infinitif est considéré comme faible en français et en allemand.

Ainsi, pour déterminer la vitalité d'une construction, Rémi-Giraud se fonde sur la complexité d'oppositions formelles. D'autres chercheurs, au contraire, considèrent les occurrences d'une forme comme facteur déterminant sa vitalité. Ainsi Simone et Cerbasi (2001, 447) constatent que « Latin [...] *is* able to see facts causatively, and, accordingly, *does* feature several types of CCs [causative constructions], but nonetheless exhibits a surprisingly *low frequency* of CCs. » Et plus tard (p. 450) : « it was a relatively poor causative-oriented language. » Alors, en dépit de l'existence d'un certain nombre de constructions causatives, les auteurs considèrent le latin comme « a poor causative-oriented language ».

Nous pensons qu'en effet c'est la fréquence de différentes constructions dans les textes, telles les constructions causatives du type *faire*/*mandar* + inf, qui est le critère de base déterminant leur vitalité. À l'occasion on peut rappeler que les vieux textes français offraient une richesse des mots. Cependant une telle richesse est illusoire tant qu'elle n'est pas fondée sur la fréquence des mots et sur la richesse des sens qu'ils expriment. Nous partons donc de la conception que la richesse des constructions se manifeste, en premier lieu, par leur fréquence dans les textes ainsi que, en deuxième lieu, par la richesse des fonctions syntaxiques qu'ils assument.

En ce qui concerne la fréquence, il convient d'ajouter les données relevées dans A et E que l'on peut comparer avec celles réunies au tableau 2 : 15 constructions du type *make* + inf dans A et 16 dans E. Ainsi, les constructions françaises sont deux fois plus nombreuses dans E et trois fois dans A. On s'aperçoit que l'original anglais de A ne favorise pas la quantité des exemples anglais.

On peut conclure que, compte tenu des deux facteurs : fréquence dans les textes et diversité des fonctions syntaxiques, la construction anglaise *make* + inf s'avère plus faible que ne le laisse suggérer la tendance analytique de cette langue.

En somme, l'infinitif anglais, y compris la construction du type *make* + inf, est relativement faible. Cette faiblesse est à rechercher dans la concurrence des formes en *-ing*, non finies elles aussi. La faiblesse de la construction causative en *make* est un facteur qui plaide en faveur de la faiblesse de l'expression causative de l'anglais. Cependant étant donné la force de ses causatifs lexicaux, on ne peut pas dire qu'il soit « a poor causative-oriented language ».

6.5. Le polonais

La faiblesse de la tendance analytique explique l'absence de la construction causative analytique en polonais. Le fait que différentes langues slaves, y compris le polonais, ont employé d'une façon limitée cette construction n'a rien d'irrégulier non plus.

6.6. Coefficient de similitude

Nous avons signalé à plusieurs reprises la plus grande richesse des constructions causatives françaises en *faire* mais aussi une position privilégiée de l'italien. Le coefficient de similitude rend compte de ce fait. Il est la quantité des exemples réunis au tableau 1 dans lesquels deux langues comparées (le français et l'espagnol, le française et l'italien, etc.) comportent un causatif en *faire*. Elle est de 29 pour le français et l'italien. Cette quantité paraît basse: (29 sur un total de 75; cf. les remarques de Iliescu (1995, 12) sur les traductions italiennes des causatifs français employés dans un livre de Simenon). Cela tient à la façon de compter les exemples. Nous retenons les exemples même dans le cas où une seule langue, pas nécessairement celle de l'original, comporte une construction en *faire*, tandis que les autres langues offrent des équivalents tels que causatif lexical, construction avec un verbe non causatif ou causatif mais non grammaticalisé (ex. *ordonner* au lieu de *faire*).

Ce coefficient est presque identique pour le français comparé à l'espagnol et au portugais: respectivement de 18 et de 17 exemples. Si l'on prenait en compte la construction causative en *mandar*, le coefficient pour le français et le portugais s'élèverait à 20. La similitude entre le français et le roumain est relativement grande (15) à ceci près que le roumain emploie le subjonctif au lieu de l'infinitif. La similitude entre le français et l'allemand s'élève à 10 exemples comportant la construction analytique,

celle entre le français et l'anglais à 7. Ces données permettent de constater que 1° la similitude entre le français et l'italien est la plus grande et que 2° la similitude entre le français et les langues romanes est plus grande que celle entre le français et les langues non romanes. Ici c'est le critère génétique qui prédomine : les langues romanes du centre, le français et l'italien, s'avèrent les plus rapprochées, tandis que les langues non romanes—les plus éloignées du français. La faiblesse du critère de similitude est qu'il montre un rapport étroit avec la fidélité des traductions, la plus facile à atteindre entre le français et une autre langue romane.

7. Remarques finales

L'article comporte deux éléments essentiels : 1° établissement de la vitalité des constructions causatives dans les cinq langues romanes les plus connues, deux langues germaniques et une langue slave sur la base d'un corpus de textes parallèles et 2° explication typologique des différences observées.

La vitalité des constructions causatives est la plus grande en français, un peu moins grande en italien. La catégorie suivante est formée par les constructions des autres langues romanes et de l'allemand. Les constructions anglaises sont plus pauvres. Le polonais de nos jours n'a pas de construction causative analytique.

Avant de proposer une explication typologique des différences observées, il convient de rappeler les catégories qui se dégagent de notre classification des langues prises en compte (Gawełko 2001, 40) : 1° l'anglais et le français, 2° les autres langues romanes, 3° l'allemand et 4° le polonais. Ces catégories vont des langues les plus analytiques vers les plus synthétiques. Les langues appartenant à la deuxième catégorie se différencient aussi par le degré plus ou moins élevé d'analytisme, mais les différences qu'elles offrent ne vont pas loin et l'influence qu'elles exercent sur la spécificité des catégories grammaticales est limitée.

Le principe général est que la construction causative analytique augmente son importance avec l'augmentation du caractère analytique de la langue où elle est employée. Les quatre catégories rappelées ci-dessus expliquent le mieux le comportement 1° du français : la grande richesse des constructions causatives en *faire* est en rapport avec le haut degré de son caractère analytique et 2° du polonais : le haut degré de synthétisme est en rapport avec l'absence de construction causative. Le comportement des langues ibéro-romanes et du roumain est régulier dans ce sens

que leurs constructions causatives sont plus faibles que celles en français, mais plus riches que celles en polonais.

L'irrégularité la plus importante est offerte par l'anglais: le haut degré d'analytisme n'est pas en accord avec la faiblesse relative de ses constructions causatives. La faiblesse en question est en rapport avec la faiblesse de son infinitif, concurrencé par les formes en *-ing*.

Le comportement de l'allemand n'est pas tout à fait régulier: il est plus synthétique que n'importe quelle langue romane et pourtant la vitalité de ses constructions causatives se rapproche de celle des constructions de la majorité des langues romanes. Nous avons proposé une explication partielle de ce fait en constatant que la différence dans le degré de synthétisme/analytisme entre la majorité des langues romanes et l'allemand ne va pas très loin.

L'irrégularité présentée par l'italien consiste en ce que son caractère typologique se rapproche de celui des langues ibéro-romanes (cf. Gawełko 2001, 40) et pourtant ses constructions causatives — des constructions françaises. Pour jeter quelque lumière sur ce fait, il convient de constater que les langues appartenant à une famille présentent parfois un comportement similaire en dépit de différences dans le degré de synthétisme/analytisme qui les caractérise. Telle la vitalité de l'infinitif, similaire dans les langues romanes, le roumain excepté. Or, en principe cette vitalité tient aussi au degré d'analytisme de la langue où il est employé. Cependant c'est l'infinitif portugais qui est un peu plus fort que l'infinitif des autres langues romanes.

Les langues ibéro-romanes et le roumain ne montrent pas d'irrégularité. Cependant chaque langue présente une certaine spécificité.

Le roumain offre un cas particulier: l'élimination progressive de l'infinitif, laquelle est un trait balkanique, entraîne la construction causative du type *face* + subjonctif.

La faiblesse relative de la construction portugaise *fazer* + inf peut étonner vu que l'infinitif portugais est le plus fort dans la Romania. Cette faiblesse relative est en rapport avec le caractère très verbal de l'infinitif portugais: sujet syntaxique relativement fréquent, p. ex. *o tempo de* ***ela*** *partir para Lisboa* (ML, 86); présence du sujet flexionnel, très rare dans les langues du monde, p. ex. *o amigo ficaria contente por* ***chegarmos*** *tão cedo* (E, 82); substantivation relativement rare (cf. Disterheft 1980, 198; Gawełko 2004).

Corpus

1° Camus, A. *L'étranger*, Paris, Gallimard, 1942; *El extranjero*, Madrid, El Libro de Bolsillo, 1982; *Lo straniero*, Milan, Tascabili Bompiani, 1988; *O estrangeiro*, Lisbonne, Unibolso, s.d.; *Străinul*, in *Străinul, Ciuma, Căderea, Exilul şi împărăţia*, Bucarest, RAO International Publishing Company, 27–92, 1993; *The Outsider*, London, Penguin Books, 1983; *Der Fremde*, Reinbeck bei Hamburg, Rowohlt, 1993; *Obcy*, Varsovie, Krąg, 1991. (E)

2° Camus, A. *La chute*, Paris, Gallimard, 1989; *La caída*, Madrid, Alianza Editorial, 1982; *La caduta*, Milan, Tascabili Bompiani, 1989; *A Queda*, Lisbonne, Livros do Brasil e Editorial Verbo, 1971; *Căderea*, in *Străinul, Ciuma, Căderea, Exilul şi împărăţia*, Bucarest, RAO International Publishing Company, 295–369, 1993; *The Fall*, London, Penguin Books, 1963; *Der Fall*, Reinbeck bei Hamburg, Rowohlt, 1995; *Upadek*, Varsovie, Krąg, 1991. (CH)

3° Camus, A. *La peste*, Paris, Gallimard, 1989; *A peste*, Lisbonne, Livros do Brasil, s.d.; *La peste*, Barcelone, Edhasa, 1977; *La peste*, Milan, Tascabili Bompiani, 1989; *Ciuma*, in *Străinul, Ciuma, Căderea, Exilul şi împărăţia*, Bucarest, RAO International Publishing Company, 93–294, 1993; *The plague*, London, Penguin Books, 1960; *Die Pest*, Hambourg, Rowohlt, 1994; *Dżuma*, Varsovie, Jota, 1991. (P)

4° Maupassant, G. de. *Une vie*, Paris, Gallimard, 1974; *Una vida*, in G. de Maupassant, *Obras completas*, 149–285, Madrid, Aguilar, 1965; *Uma vida*, Mem Martins, Livros de bolso Europa-America, 1974; *Una vita*, Farigliano, Arnoldo Mondadori Editore, 1984; *O viaţă*, Chişinău, Făt-Frumos, 1994; *A Woman's Life*, London, Penguin Books, 1965; *Ein Leben*, Frankfurt/Main & Berlin, Ullstein, 1995; *Historia jednego życia*, Varsovie, Zrzeszenie Księgarstwa, 1985. (V)

5° *Memórias da Irmã Lúcia*, Fátima, Postulação, 1978; *Mémoires de Sœur Lucie*, Fátima, Postulação, 1980; *Memorias de la Hermana Lucía*, Fátima, Postulação, 1978; *Memorie di Suor Lucia*, Fátima, Postulação, 1980; *Sora Lucia despre Fatima*, Fátima, Secretariado dos Pastorinhos, 2002; *Fatima in Lucia's own words*, Fátima, Postulação, 1976; *Schwester Lucia spricht über Fátima*, Fátima, 1975; *Siostra Łucja mówi o Fatimie*, Fátima, Postulação, 1978 (jusqu'à la p. 104 de l'original portugais). (ML)

6° Tomas a Kempis, *De imitatione Christi*; (*O naśladowaniu Jezusa Chrystusa*), édition bilingue, Varsovie, Instytut Wydawniczy Pax, 1981; *L'imitation de Jésus-Christ*, Paris, Édition du Seuil, 1961; *L'imitazione di Cristo*, Milan, Edizioni Paoline, 1988; *Imitación de Cristo*, Barcelone, Editorial Regina, 1987; *Imitação de Cristo*, Braga, Editorial Franciscana, 1980; *Urmarea lui Hristos*, Timişoara, Editura Metropoliei Banatului, 1991; *The Imitation of Christ*, London, Penguin Books, 1965; Thomas von Kempen, *Die Nachfolge Christi*, Kevelaer, Butzon & Bercker, 1990; Tomasz à Kempis, *Naśladowanie Chrystusa*, Cracovie, WAM, 2000. (IC)

7° Carrol, L. *Alice's adventures in Wonderland — Les aventures d'Alice au Pays des merveilles*, Paris, Aubier-Flammarion, 1970; *Aventuras de Alicia en el país de las maravillas*, Barcelone, Moby Dick, 1973; *Alice nel Paese delle Meraviglie*, in *Alice nel Paese delle Meraviglie, Attraverso lo specchio*, Milan, Garzanti, 1–141, 1989; *Alice no País das Maravilhas*, Mem Martins, Publicações Europa-America, 1998; *Alisa în ţara minunilor*, Craiova, Editura Literatorul, 1991; *Alice im Wunderland*, Hamburg, Cecilie Dressler Verlag, 1990; *Przygody Alicji w Krainie Czarów*, Varsovie, Lettrex, 1990. (A)

8° Sienkiewicz, H. *Quo Vadis*, Varsovie, Labos, 1990; *Quo vadis*, Paris, Le livre de poche 3161, 1971; *Quo vadis ?* Madrid, Aguilar, 1967; *Quo vadis ?*, Milan, Biblioteca

Universale Rizzoli, 1984; *Quo Vadis ?*, Mem Martins, Livros de bolso Europa-America, 1974; *Quo Vadis*, Bucarest, Editura Universul, 1945; *Quo Vadis ?*, Sawtry-New York, Hippocrene-Dedalus, 1993; *Quo Vadis*, Zurich, Diogenes, 1985 (chap. X–XII du premier volume). (Q)

9° Andrzejewski, J. *Popiół i diament*, Varsovie, Krajowa Agencja Wydawnicza, 1966; *Cendres et diamant*, Paris, Gallimard, 1967; *Cenizas y diamantes*, Barcelone, Luis de Caralt, 1966; *Cenere e diamanti*, Lerici Editori, Milan, 1961; *Cenuşă şi diamant*, Bucarest, Editura pentru Literatură Universală, 1968; *Ashes and Diamond*, Eveston, Northwestern Univ. Press, 1991; *Asche und Diamant*, Frankfurt/Main, Suhrkamp Verlag, 1984 (chap. IV et V) (CD)

10° Maupassant, G. de. *Bel ami*, Livros de bolso Europa-América, Guimarães Editores, 1972. (BA)

Bibliographie

Aranda Ortiz, Antonio 1990. La expresión de la causatividad en español actual. Libros Pórtico, Zaragoza.

Comrie, Bernard 1976. The syntax of causative constructions: Cross-language similarities and divergencies. In: Shibatani (1976b, 261–312).

Coşeriu, Eugenio 1988. Der romanische Sprachtypus. Versuch einer neuen Typologisierung der romanischen Sprachen. In: Energeia und Ergon 1 : 207–24.

Disterheft, Dorothy 1980. The syntactic development of the infinitive in Indo-European. Slavica Publishers, Columbus, Ohio.

Duffley, Patrick 1992. The English infinitive. Longman, London.

Gawełko, Marek 2001. Essai de classification fonctionnelle des langues romanes. In: Romance Philology 55 : 21–40.

Gawełko, Marek 2004. Quelques particularités de l'infinitif substantivé en espagnol et en portugais. In: Bulletin Hispanique 2 : 615–35.

Givón, Talmy – Phil Young 2002. Cooperation and interpersonal manipulation in the society of intimates. In: Shibatani (2002a, 23–56).

Gołąb, Zbigniew 1968. The grammar of Slavic causatives. In: Henry Kučera (ed.): Amaerican contributions to the Sixth International Congress of Slavists, vol. 1, 71–94. Mouton, Prague.

Heinisch, R. 1977. Vergleichende Untersuchungen zu slovakischen Präfixverben und ihren deutschen Entsprechungen. Leipzig (inédit).

Hernanz Carbó, María Luisa 2000. El infinitivo. In: Ignecio Bosque – Violeta Demonte (eds): Gramática descriptiva de la lengua española, 2197–356. Espasa, Madrid.

Iliescu, Maria 1993. Le factitif roumain. In: Revue Roumaine de Linguistique 38 : 297–305.

Iliescu, Maria 1995. La factitif dans les langues romanes. In: Romanistik in Geschichte und Gegenwart 2 : 1–14.

Marstrander, Leiv-Otto 2000. La proposition infinitive et le passif. In: Lene Schøsler (ed.): Actes du colloque international, Institut d'Études Romanes, Université de Copenhague, 5–7 mars 1998, 205–211. Museum Tusculanum Press, Copenhagen.

Moeschler, Jacques 2003. L'expression de la causalité en français. In : Cahiers de Linguistique Française 25 : 11–42.

Mori, Olga 1980. Frases infinitivas preposicionales en la zona significativa causal. Gunter Narr Verlag, Tübingen.

Pobożniak, Tadeusz 1979. Die Causativa in der Hindi-Sprache. In : Folia Orientalia 20 : 5–21.

Reboul, Anne 2003. Causalité, force dynamique et ramifications temporelles. In : Cahiers de linguistique française 25 : 43–69.

Rémi-Giraud, Sylvianne 1988. Les grilles de Procuste, Description comparée de l'infinitif en français, grec ancien, allemand, anglais et arabe. In : Sylvianne Rémi-Giraud (ed.) : L'infinitif. Une approche comparative, 11–68. Presse universitaire de Lyon, Lyon.

Robustelli, Cecilia 1995. La costruzione di Fare con l'Infinito in Italiano moderno. In : Studi e Saggi linguistici 25 : 199–272.

Sandfeld, Kristian 1965. Syntaxe du français contemporain : L'infinitif. Librairie Droz, Genève. (1ère éd. 1943).

Shibatani, Masayoshi 1976a. The grammar of causative constructions : A conspectus. In : Shibatani (1976b, 1–40).

Shibatani, Masayoshi (ed.) 1976b. The grammar of causative constructions (Syntax and semantics, vol. 6). John Benjamins, Amsterdam & Philadelphia.

Shibatani, Masayoshi (ed.) 2002a. The grammar of causation and interpersonal manipulation. Academic Press, New York.

Shibatani, Masayoshi 2002b. Introduction : Some basic issues in the grammar of causation. In : Shibatani (2002a, 1–22).

Shibatani, Masayoshi – Prashant Pardeshi 2002. The causative continuum. In : Shibatani (2002a, 85–126).

Simone, Raffaele – Donato Cerbasi 2001. Types and diachronic evolution of Romance causative constructions. In : Romanische Forschungen 113 : 441–473.

Termińska, Kamila 1983. Składnia czasowników kauzatywnych we współczesnym języku polskim. Uniwersytet Śląski, Katowice.

Urbańczyk, Stanisław – Marian Kucała 1999. Encyklopedia języka polskiego, 3e éd. Ossolineum, Wrocław.

Vanvolsem, Serge 1983. L'infinito sostantivato in italiano. Accademia della Crusca, Firenze.

Remarks on *make* + infinitive type causative constructions in Romance and some other languages

This paper establishes the productivity of causative constructions in eight languages, two of which (English and French) are highly analytical, one (Polish) highly synthetic, and five show a moderate tendency to analyticity. The paper then proceeds to a typological explanation of the differences observed. The general conclusion is that the productivity of causative constructions is proportionate to the degree of analyticity of the language : it is the strongest in French and the weakest in Polish. English is an exception in that causative constructions occur relatively rarely in it, due to the relative infrequency of the infinitive.

Keywords : causative, productivity, explanation, synthetic, analytical

Acta Linguistica Hungarica, Vol. 53 (2), pp. 139–174 (2006)
DOI: 10.1556/ALing.53.2006.2.3

TOPIC, LOGICAL SUBJECT AND SENTENCE STRUCTURE IN HUNGARIAN

ZSUZSANNA GÉCSEG

Department of French Language and Literature
University of Szeged
Egyetem utca 2.
H–6722 Szeged
Hungary
gecsegz@lit.u-szeged.hu

Abstract: This paper investigates the "discourse-configurationality" hypothesis in Hungarian, based on the current assumption that Hungarian sentence structure is largely determined by information structure. It argues for the necessity of differentiating between the notion of topic, defined on a pragmatic level with respect to possible contexts, and the notion of logical subject, defined on a decontextualized logico-semantic level. On the basis of the distinction between these two levels of sentence analysis, Hungarian should be taken as a logical subject-prominent language rather than a topic-prominent one. As for the so-called contrastive topic in Hungarian, only a subclass of contrastive topic expressions meets the topicality conditions established in this paper on a pragmatic ground, and other types of contrastive topic expressions, namely those that can hardly be differentiated from ordinary topics, display the properties of logical subjects rather than topics.

Keywords: topic, logical subject, word order, thetic judgment, categorical judgment

Introduction

The aim of this paper is to investigate the so-called "discourse-configurationality" of Hungarian, based on the current assumption that Hungarian sentence structure is largely determined by information structure. According to this assumption, Hungarian sentences belong to two basic types: the first has a topic–comment (or logical subject–logical predicate) structure and the second is topicless, consisting of a complex logical predicate (see É. Kiss 1994; 2002 and Szabolcsi 1997). The topic constituents

1216–8076/$ 20.00

occupy the first position(s) in the preverbal field (the Spec of one or several TopPs dominating each other) and the comment is constituted by the VP and various operators (such as distributive quantifiers, focalized expressions and verbal modifiers) preceding the verb. The two types of sentence structure are exemplified by (1) and (2), respectively:

(1) Marit elütötte a vonat.
Mary-acc ran over the train
'Mary has been run over by the train.'

(2) (a) Megérkezett Feri.
arrived Feri
'Feri has arrived.'

(b) Minden gyerek/ (még) Feri is/ sok gyerek olvasta ezt a könyvet.
all/every child/ (even) Feri also/ many child read this the book
'Every child/even Feri has read this book./
Many children have read this book'

In (1), the constituent *Marit* 'Mary-acc' plays the role of the topic; i.e., it refers to the individual about which something is asserted by the predicate *elütötte a vonat* 'has been run over by the train'. As for (2a), the verb *megérkezett* 'has arrived' occupies a sentence-initial position and the sentence is not interpreted as an assertion about Feri, but as a complex predicate reporting a certain event. In (2b), the preverbal positions are filled by distributive quantifiers that do not serve to denote an individual, but to quantify over individuals, consequently, they take part in the main predication expressed by the whole sentence.

In this paper I will argue for the claim that the terms "topic" and "logical subject" are far from being synonymous, but belong to two distinct levels of information structure: the former is contextually determined and as such is part of the so-called pragmatic articulation, whereas the latter should be taken as a context-free notion, defined on a semantico-logical ground.

After an overview of current assumptions in Hungarian linguistics about the properties of the logical subject in section **1**, problems connected with these assumptions will be raised in section **2**. In section **3**, it will be pointed out that the distinction between thetic judgments and categorical judgments, often related to the distinction between topicless sentences and those with a topic–comment articulation, has been interpreted in rather divergent ways since the works of Marty and Brentano

and these divergences can be the source of confusion between separate levels of sentence analysis. In section 4, I will present my proposal, namely that word order in Hungarian sentences is determined on a semantico-logical level of information structure rather than a pragmatic one. The proposal will be based on two distinct definitions of the notions of logical subject and topic. As a consequence, a constituent that plays the role of logical subject is not necessarily the topic of the sentence in a particular context. I will also argue for the claim that contrastive topics display some properties that show that these constituents are more related to pragmatic articulation than the ordinary topic is. Finally, in section 5, I will briefly characterize a spoken version of French, called colloquial French, and I will point out that this version of French is essentially a topic-prominent language, as opposed to Hungarian, which should be taken as having a logical subject–logical predicate articulation.

1. The properties of topic (logical subject) expressions in Hungarian

In Hungarian, the constituents[1] that play the role of topic in a particular sentence are usually characterized by means of syntactic, prosodic and semantic criteria.

If the preverbal field of the sentence is filled by some material, its constituents must appear in a strict order. The sentence topic occupies the leftmost position, followed by the logical predicate, consisting of different kinds of operators (distributive quantifiers, focalized expressions) or a verbal modifier as indicated by the scheme (3) (É. Kiss 1994; 1998; Szabolcsi 1997):

(3) Topic* – Distributive Quantifier* – Focus/Verbal modifier – Verb – Postverbal Material

The boundary between the topic and the logical predicate can usually be identified by prosodic means[2] as well: according to É. Kiss (2002, 11),

[1] The main purpose of this paper is to investigate the discourse function of nominal expressions (called NPs for the sake of simplicity); therefore, other categories (such as adverbials) that can play the role of topic will be neglected here.

[2] On the difficulties of applying the prosodic criterion in the case of a syntactically complex topic, see Kálmán (2001).

"the first obligatory stress, which also represents the heaviest grammatical stress in the sentence, falls on the first major constituent of the predicate."

As for the discourse function of topic, it is defined as denoting the individual about which something is asserted by the logical predicate (É. Kiss 1994; Kálmán 2001). Since the referent of a topic expression must always be identifiable in order to be able to assert something about it (É. Kiss 2000; Maleczki 2003), the topic expression must be a referring expression, i.e., it must be referential and specific.[3] The term "specificity" should be taken in the sense of Enç (1991), who considers an NP specific if it refers to an individual anchored in the discourse context; that is, the referent of a specific NP is either already introduced in the discourse or is a member of a contextually determined set. These requirements are reflected by the following definition proposed by É. Kiss:

> "The topic foregrounds an individual (a person, an object, or a group of them) from among those present in the universe of discourse as the subject of the subsequent predication." (É. Kiss 2002, 9)

It follows from these properties that among non-generic expressions only definite NPs (proper names such as *Péter*, *Mari* or definite descriptions such as *a fiú* 'the boy', *ezek a lányok* 'these girls' etc.) and specific indefinite NPs (indefinite pronouns such as *valaki* 'somebody', common nouns introduced by numerals such as *két gyerek* 'two (of the) children' or indefinite determiners such as *valamelyik gyerek* 'one of the children') can occupy the position of the topic of the sentence.[4]

On the basis of these claims, we can characterize sentences (4)–(6) as having a topic–comment (logical subject–logical predicate) articulation:

(4) Mari `szereti az almás pitét.
Mary likes the apple pie
'Mary likes apple pie.'

[3] As É. Kiss (2002) points out, generic NPs also meet the requirement of identifiability—in this sense they should be taken as referential expressions; that is, they can play the role of the sentence topic.

[4] There is a certain disagreement concerning the information structure status of NPs introduced by *a legtöbb* 'most': Szabolcsi (1997) considers them as topic expressions, whereas É. Kiss (1998) and Kálmán (2001) argue for their quantificational/predicative nature.

(5) Két gyerek már `megette az ebédjét.
two child already eat the lunch-his
'Two children have already finished their lunch.'

(6) Sok barátom `jobban szereti a klasszikus zenét, mint a rockot.
many friend-my better like the classical music than the rock
'Many of my friends prefer classical music to rock.'

In (4), the position of the topic is filled by the proper name *Mari*, denoting an individual present in the universe of discourse of the participants in the sense that this individual is supposed to be known by speaker and listener alike; the predicate asserts that the individual called Mari likes apple pie.

Sentence (5) contains a specific indefinite NP (*két gyerek* 'two children') playing the role of the topic. This NP denotes a subgroup of a contextually determined set of children and the predicate asserts about this subgroup that its members have already finished their lunch.

As for (6), the position of the topic is occupied here by the indefinite *sok barátom* 'many of my friends' denoting a relatively large subclass of a group of people fully identifiable for the speaker. The predicate of the sentence asserts about this subclass that its members prefer classical music to rock.

The topic–comment articulation of (4)–(6) can be supported by prosodic arguments: although they can be pronounced in several ways, in each of them the constituent marked by "`" bears the first obligatory accent of the string.

Sentences (7)–(8) below illustrate the sentence type that is topicless and consists of a complex predicate covering the whole sentence:

(7) `Sok barátom eljött a partira.
many friend-my came the party-to
'Many of my friends came to the party.'

(8) `Minden barátom/ még `Péter is szereti a klasszikus zenét.
all friend-my/ yet Peter also likes the classical music
'All of my friends like classical music./
Even Peter likes classical music.'

Both sentences contain a distributive quantifier in the sentence-initial position bearing the first obligatory accent of the sentence. The position

of the topic remains unfilled in these sentences; consequently, they lack topic–comment articulation.

There is a special kind of topic expression, called contrastive topic, which occurs in the same syntactic position and is claimed to have the same information structure status as the ordinary topic; at the same time, contrastive topics have certain prosodic and semantic features distinguishing them from other kinds of topics.[5]

First, contrastive topics are characterized by a prosodic prominence and a fall-rise intonation contour (noted by "ˇ" below) and by the obligatory presence of another prosodically prominent element of the sentence called the associate of the contrastive topic:

(9) ˇMari az `almás pitét szereti.
Mary the apple pie likes
'As for Mary, she likes apple pie.'

In (9), *Mari* is pronounced with a (fall)-rise intonation contour, and the constituent *az almás pitét* 'apple pie' has focus stress.

Second, a sentence containing a contrastive topic is claimed to bear a special implicature, related to the fact that the contrastive topic constituent refers to an individual being a member of a set consisting of the alternatives of this individual. Such a sentence implies that there are other individuals in the set of alternatives for which the main predicate of the sentence (possibly) does not hold. In the case of (9), for instance, the individual denoted by *Mari* is contrasted with other individuals in the discourse domain being possibly fond of dishes other than apple pie.

In Hungarian, it is not only referential expressions that can function as contrastive topic, but other types of expressions (such as certain quantifiers or bare common nouns) that otherwise would never occur in the position of the sentence topic. As É. Kiss (2000) and É. Kiss–Gyuris (2003) point out, this can be explained by the fact that an expression functioning as a contrastive topic can denote a higher order individual, for instance (the name of) a property in the case of a bare common noun or a property of plural individuals, in the case of a quantifier. This also explains the apparent scope inversion attested in sentences containing a quantifier playing the role of the contrastive topic:

[5] For an overview of the diverse conceptions of the topical and focal character of contrastive topics, see Molnár (1998). For a detailed analysis of contrastive topic in Hungarian, see Gyuris (2002).

(10) ˇMarslakót még `soha nem láttam.
Martian yet never no saw-1sg
'I have never seen a Martian.'

(11) ˇLegalább két filmet `mindegyik diák megnézett.
at least two film-acc each student watched
'Each student saw at least two films.'

Sentence (10) asserts about the property of being a Martian that the speaker has never seen an individual with this property. On the other hand, the sentence implies that there are possibly other properties comparable to that of being a Martian (any kind of strange property), instances of which the speaker has already seen. As for (11), the sentence asserts about the property of being a plurality consisting of two films that each student saw an instance of this property.

2. Some problems with assumptions about the notion of topic

2.1. Specificity

In order to be able to appear in the position of topic, an NP in Hungarian must meet the condition of specificity. However, as Maleczki (2001; 2003) points out, this notion cannot be considered a dichotomy; rather, there are expressions that are more or less identifiable in a given context.

An expression can be fully identified if its referent is part of the situation or if its referent is known by both speaker and listener—this is the case of proper names, deictic pronouns or definite descriptions related to speaker or listener (*my mother, your friends*, etc.). Definite descriptions can also be identified by means of discourse context (a definite description normally refers to an individual already mentioned in the preceding context). Finally, an indefinite is (more or less) specific if its referent is somehow linked to the situation or to the preceding context. Consider the following examples:

(12) Egy gyerek virágot szed a kertedben.
a child flower-acc picks the garden-your-in
'A child is picking flowers in your garden.' (Maleczki 2001, 161)

(13) Egy gyerek már felébredt.
a child already pfx-woke
'A child has already woken up.' (*idem.*)

(14) Egy gyerek kiszaladt az úttestre.
a child pfx-ran the roadway-onto
'A child has run out onto the roadway.' (*ibid.*, 162)

According to Maleczki (2001), in (12) the most natural interpretation of *egy gyerek* 'a child' is a totally unknown child, in (13) the subject NP refers to a member of a previously given set of children, whereas in (14) the same NP is ambiguous with respect to specificity: its referent can be unidentified or specific. Notice, however, that in the case of (14), discourse-linkedness can be taken in several ways: the subject NP can denote a member of a previously mentioned set of children or just a member of the set of passers-by. Even in this case we can talk about a discourse-linked referent, since a discourse that describes events that take place in the street naturally presupposes the existence of certain elements of the events such as vehicles and passers-by.

Similarly, the sentence-initial indefinite NPs in (15) are difficult to analyse in terms of specificity:

(15) Egy újságban/ az egyik újságban azt olvastam,
a newspaper-in the one newspaper-in that-acc read-1sg
hogy ötösikrek születtek Új-Zélandon.
that quintuplets were born New Zealand-on
'I read in a newspaper/one of the newspapers that quintuplets were born in New Zealand.'

The difference between NPs of the type *egy újság* 'a newspaper' and *az egyik újság* 'one of the newspapers' is that the former is ambiguous between a specific and a non-specific reading, whereas the latter has an unambiguously specific (i.e., partitive) reading. However, as (15) illustrates, both can be used to refer to a particular newspaper without having previously mentioned a set of newspapers. One can argue that we do not need any previously introduced superset here since the sentence-initial NP denotes a member of a maximal set, i.e., the set of all (relevant) newspapers—this maximal set is comparable to the referent of a generic NP and as such, is identifiable independently of any context. The problem is, however, that along the lines of such reasoning any preverbal (non-

bare) indefinite NP should be considered specific, including the subject of (12) above.

Lambrecht (1994) makes a distinction between a specific and a non-specific NP by claiming that the referent of the former is identifiable for the speaker but not identifiable for the listener, while the referent of the latter is not identifiable for the speaker either. However, the discourse-linkedness of a non-specific NP does not necessarily guarantee its identifiability. Consider (16)–(17):

(16) Az egyik gyerek hiányzik.
the one child misses
'One of the children is missing.'

(17) Valaki/ Valamelyik diák elvesztette a laptopját.
somebody some student lost the laptop-his-acc
'Somebody/one of the students has lost his laptop.'

One can utter (16) without being able to identify the referent of the NP *az egyik gyerek* 'one of the children'—it is enough to count the children in question; that is, one can identify a set without being able to identify its members taken individually. Similarly, we don't have to identify the student in question in order to be able to utter (17)—the sentence can be uttered in a situation where the speaker sees an abandoned laptop in an empty lecture hall. On the other hand, one can indifferently use *valaki* 'somebody' instead of *valamelyik diák* 'one of the students' without perceiving any difference in their information structure status or syntactic behavior.

É. Kiss (2002, 11) is also aware of the fact that NPs denoting non-identified (i.e., non-specific) individuals can appear in topic position. She exemplifies this with (18a–b) below:

(18) (a) Valaki kopog.
somebody knocks
'Somebody is knocking.'

(b) Valami leesett a tetőről.
something pfx-fell the roof-from
'Something has fallen from the roof.'

É. Kiss proposes to consider here the NPs *valaki* 'somebody' and *valami* 'something' as specific in the sense that (18a–b) are normally used "in

situations in which the existence of an unidentified person or object has been inferred—e.g., when knocking has been heard at the door, or an object has been seen flying past the window, respectively" (É. Kiss 2002, 11). However, this reasoning is weakened by the existence of examples like (19):

(19) Valaki tegnap bekopogott az ajtómon.
somebody yesterday pfx-knocked the door-my-on
'Somebody knocked at my door yesterday.'

In fact, at the moment of the utterance of (19) there is obviously nothing that allows one to infer the existence of the referent of *valaki*—sentence (19) asserts the existence of an individual with the property of having knocked at the speaker's door the day before.[6]

All these arguments tend to the conclusion that NPs that denote completely unidentified referents can also appear in topic position. On the basis of the assumption that an indefinite that denotes a brand-new referent cannot function as the topic of the sentence, Maleczki (2003) claims that such an NP—even if it appears in the preverbal field—does not fill the position of the sentence topic. However, one cannot postulate an intermediate syntactic position between the topic and distributive quantifiers, since non-specific indefinites can even precede specific NPs in the preverbal field:

(20) (a) Marinak valaki szerelmesleveleket ír.
Mary-to somebody love letters-acc writes
(b) ?Valaki Marinak szerelmesleveleket ír.
somebody Mary-to love letters-acc writes
'Somebody writes love letters to Mary.'

In fact, the version in which a specific (for example a definite) NP precedes a non-specific one (as in (20a)) is slightly more acceptable than the opposite word order (exemplified in (20b)), but this can be explained by a general requirement according to which an NP higher on the specificity scale tends to precede an NP lower on that scale:

[6] On the other hand, (19) illustrates the application of a topic-test frequently used in Hungarian syntax. Unstressed sentence adverbials such as *tegnap* 'yesterday' are claimed to precede or follow the topic but never be part of the predicate. According to this test, the NP *valaki* 'somebody' fills the topic position here.

(21) (a) Marinak valamelyik diák szerelmesleveleket ír.
Mary-to some student love letters-acc writes

(b) [7]Valamelyik diák Marinak szerelmesleveleket ír.
some student Mary-to love letters-acc writes
'One of the students writes love letters to Mary.'

Although (21a–b) contain only specific NPs in topic positions, (21a) is more acceptable since a definite NP is more specific than a specific indefinite NP.[7]

2.2. Ambiguity with respect to the thetic–categorical distinction

Following a terminology originally used by Marty (1918) and Brentano (1973), a sentence with a topic–comment articulation is often called a categorical sentence (i.e., expressing a categorical judgment), whereas a topicless sentence is called a thetic sentence (i.e., expressing a thetic judgment). The difference between the two sentence types can be captured by means of preceding questions for which the sentence under consideration constitutes an adequate answer. Since a categorical sentence asserts something about a topic referent, this sentence type normally answers a *wh*-question about this referent. On the other hand, thetic sentences are claimed to assert something about a particular event or situation; hence they are appropriate answers to questions of the type "What happened?", "What's the news?", "What's the problem?".

The question–answer pairs of the two types described above are exemplified in (22) and (23):

(22) Q: Hol van Mari?
'Where is Mary?'

A: Mari a kertben dolgozik.
Mary the garden-in works
'Mary is working in the garden.'

[7] The same prominence requirement can be formulated in terms of subject prominence as well as animate prominence: in fact, a subject tends to precede a non-subject and an NP denoting an animate referent tends to precede an NP denoting an inanimate in topic position. These requirements are sometimes in conflict with each other, resulting in more or less acceptable sentences.

(23) Q: Mi történt?
'What happened?'

A: Megérkeztek a vendégek.
arrived-pfx the guests
'The guests have arrived.'

In (22), sentence A expresses a categorical judgment in the sense that it asserts about Mary that she is working in the garden. On the other hand, sentence A in (23) corresponds to a thetic judgment since it asserts about a particular situation that the arrival of the guests took place in that situation.

In Hungarian, a sentence in which the topic position is filled by a constituent is claimed to express a categorical judgment whereas a sentence with a verb in sentence-initial position is considered to express a thetic judgment. However, this is only partially true. In fact, a sentence with a verb in the leftmost position can never answer a question about the referent of a constituent in its postverbal domain; that is, question–answer pairs of the type (24) are never acceptable:

(24) Q: Mi történt Gézával?
'What's happened to Géza?'

A: #Meghalt Géza.[8]
died-pfx Géza
'Géza is dead.'

However, a sentence in which the topic position is filled can often (though not always) function as a correct answer to the question "What's happened?". Consider for instance (25):

(25) Mari beteg/ megbetegedett.
Mary ill became-ill-pfx
'Mary is ill/has got ill.'

Sentence (25) can constitute a correct answer to both types of question, i.e., to "What's happened?" as well as to "What's happened to Mary?". Moreover, in this particular case, the verb-initial version would be odd even if it replies to the question "What's happened?":

[8] The symbol "#" marks that the sentence under consideration is acceptable but does not fit the preceding question

(26) Q: Mi történt?
'What happened?'

A: *?Beteg/ Megbetegedett Mari.
ill became-ill-pfx Mary
'Mary is ill/has got ill.'

Since (25) can appear in two types of contexts exemplified by the two types of questions, we conclude that the sentence is ambiguous with respect to the thetic–categorical distinction. If we identify theticity with the property of lacking a topic constituent, we must claim that in its thetic reading the constituent that fills the topic position in (25) does not play the role of the topic of the sentence.

Even though there is no structural difference related to the two readings of sentences like (25), they can be uttered following two types of intonation pattern: in the first, the constituent in topic position is unaccented, while in the second accent is equally distributed to each major constituent of the sentence, including the one occupying the topic position. As Varga (1987) points out, the latter corresponds to a reading in which the constituent in topic position introduces a brand-new referent. This brand-new referent can in fact be denoted by a specific (or even by a definite) NP if this referent has not been mentioned in the preceding discourse.

2.3. Two types of thetic sentences

We have seen in sections **2.1** and **2.2** that it is not only verb-initial sentences that can express a thetic judgment, but also sentences in which the topic position is filled with a non-specific or even a specific (definite) expression.

Moreover, as Gécseg and Kiefer (forthcoming) point out, a sentence that expresses a thetic judgment, i.e., an assertion about a particular event, can be realized by using different word orders. Consider sentences (27)–(28):

(27) (a) Bodri megharapta Marit.
Bodri bit Mary-acc
'Bodri has bitten Mary.'

(b) Marit megharapta Bodri.
Mary-acc bit Bodri
'Mary has been bitten by Bodri.'

Sentence (27a) can be uttered by means of two neutral intonation patterns corresponding to two types of contexts illustrated by the questions (28a–b):

(28) (a) Mit csinált Bodri?
'What's Bodri done?'

(b) Mi történt?
'What's happened?'

In the context created by (28a), the constituent *Bodri* is unaccented and the sentence expresses a categorical judgment about the referent of *Bodri*, while in the context created by (28b) sentence accent is equally distributed to *Bodri*, *megharapta* and *Marit* and the sentence expresses a thetic judgment about a particular event of Mary's being bitten by Bodri. Similarly, with the two neutral intonation patterns indicated above, (27b) can serve as an answer to a question about Mary as well as to a question about a particular event.

It follows from these facts that (27a–b) are synonymous in their thetic reading in the sense that they can appear in the very same context (created by the question "What's happened?"). However, in spite of the truth-conditional and contextual equivalence that characterizes the variants (a) and (b), there is an important difference between them. Independently of the possible contexts in which they can be used, the difference between their word order suggests that in (27a) something is asserted about Bodri, and in (27b) something is asserted about Mary. This can be accounted for by scheme (1) only if we consider *Bodri* and *Mari* to be the topic of sentences (27a) and (27b), respectively. However, this cannot be the case if we interpret them as thetic sentences.

Consider now (29a–b):

(29) (a) Egy kutya megharapott egy járókelőt.
a dog bit a passer-by-acc
'A dog has bitten a passer-by.'

(b) Egy járókelőt megharapott egy kutya.
a passer-by-acc bit a dog
'A passer-by has been bitten by a dog.'

Since the constituents *egy kutya* 'a dog' and *egy járókelőt* 'a passer-by-acc' denote completely unknown referents here, only the thetic (i.e., topicless) reading is available for the two word order variants.[9] Yet, as we saw in the case of (27a–b), in this reading the two sentences must be taken to be contextually equivalent.[10] However, if we admit the assumption that any change in word order must be motivated on some ground, we must assume that there is a difference in their interpretation. In fact, independently of their contextual equivalence, (29a) asserts something not only about a particular event, but also about a(n unidentified) dog and similarly, (29b) asserts something about the same event but from the point of view of another participant of this event, a(n unidentified) passer-by.

2.4. Ordinary topic vs. contrastive topic

As we saw in section **1**, the distinction between ordinary and contrastive topic is based primarily on prosodic and semantic criteria. In many cases, however, it is very difficult to decide whether a constituent in topic position should be taken to be an ordinary or contrastive topic.

Compare sentences (30a–b):

(30) (a) Mari ˋPétert csókolta meg.

(b) ˇMari ˋPétert csókolta meg.
Mary Peter-acc kissed pfx
'As for Mary, it is Peter that she kissed.'

In both sentences, the constituent *Mari* is in topic position and the sentences contain a focused expression (*Pétert*). In (30a), the topic is unaccented while in (30b) the same constituent is pronounced with a prominent stress and a rising intonation contour. Moreover, (30a–b) appear in two context types, exemplified by (31a) and (31b), respectively:

(31) (a) Kit csókolt meg Mari?
'Whom did Mary kiss?'

[9] This is also reflected by their intonation pattern: they can only be uttered with equally distributed accents on the common nouns and the (prefixed) verb in them.

[10] This confirms the assumption made by Kuroda (1972) who claims that passivization in languages like English cannot be taken to play the specific role of making the object the topic (in his terminology, the logical subject) of the sentence, since such an object can denote an indefinite (i.e., unidentified) referent as well.

(b) Kit csókoltak meg a lányok?
'Whom did the girls kiss?'

The difference between the two types of question originates in the fact that the sentence (30b) that contains a contrastive topic implies that there can be somebody else in the discourse domain who possibly kissed somebody other than Peter. Notice that this possibility is not excluded by (30a), but the intonation contour of (30b) makes this possibility more explicit, i.e., (30b) expresses that such an eventuality is "under consideration".

If we replace the definite expression in topic position with an indefinite NP[11] with a specific reading in (30), the difference between the two types of topic tends to disappear:

(32) (a) Az egyik lány ˋPétert csókolta meg.
(b) ˇAz egyik lány ˋPétert csókolta meg.
the one girl Peter-acc kissed pfx
'One of the girls kissed ˋPeter.'

(33) (a) Sok lány ˋPétert csókolta meg.
(b) ˇSok lány ˋPétert csókolta meg.
many girl Peter-acc kissed pfx
'Many of the girls kissed ˋPeter'

For sentences (32b) and (33b), if we consider only the reading in which the sentence-initial NP has a scope over the focused one (the other possibility will be accounted for later), (32a–b) and (33a–b) can be used to answer the very same question, i.e., (34):

(34) Kit csókoltak meg a lányok?
'Whom did the girls kiss?'

In fact, the difference between the two intonation contours which characterize the (a) vs. (b) sentences is very hard to perceive and the two types of sentences can often replace each other in the same context. This

[11] Although the topic NP in (32) is introduced by the definite determiner *a(z)* 'the', the NP as a whole must be considered as an indefinite expression since it denotes a member of the set (of sets) denoted by its common noun part in the same way as the indefinite *két lány* 'two girls, two of the girls' denotes a member of the set (of sets) denoted by the common noun *lány* 'girl'.

is explained by the fact that specific indefinites have *per definitionem* a partitive reading; that is, they denote a member (or a subset) of a contextually determined set. Their contrast implicature is equivalent in this case to an aspect of the Gricean maxim of quantity: the use of a partitive expression should imply that there are other relevant members of the contextually determined set for which the predicate of the sentence does not hold. Sentences (33a–b) are particularly interesting in this respect: NPs of the type *sok N* 'many Ns' can appear in topic position only if the sentence contains a constituent with prominent stress, for example a focused constituent.[12] Compare (35) and (36):

(35) (a) Mari `megérkezett Szegedre.

(b) *`Mari megérkezett Szegedre.
Mary arrived-pfx Szeged-to
'Mary has arrived in Szeged.'

(36) (a) *Sok lány `megérkezett Szegedre.

(b) `Sok lány megérkezett Szegedre.
many girl arrived-pfx Szeged-to
'Many girls have arrived in Szeged.'

Since in sentences (35)–(36) the verb *megérkezett* 'has arrived' cannot bear contrastive stress, they are never uttered with the same intonation pattern. To put it differently, the NP *sok lány* 'many girls' cannot have an (ordinary or contrastive) topic accent but must bear the first obligatory stress of the sentence, it must consequently be taken to belong to the logical predicate (i.e., comment) of the sentence.

On the other hand, a sentence containing a verb that usually has prominent stress in Hungarian such as *szeret* 'likes' allows both types of accentuation for *sok N* 'many Ns':

(37) (a) Sok gyerek `szereti az almás sütit.

(b) `Sok gyerek szereti az almás sütit.
many child likes the apple pie
'Many children like apple pie.'

[12] Gécseg (2001) characterizes NPs introduced by the determiner *a legtöbb* 'most' in a similar way: she considers NPs of this type to be inherently contrastive topics on the basis of the claim that they always appear in a sentence with a prominent associate.

In (37a) *sok gyerek* 'many children' fills the topic position of the sentence, whereas in (37b) it belongs to the logical predicate. However, (37a) appears only in very special, polemic contexts like the one exemplified by dialogue (38):

(38) A: Sok gyerek ˋgyűlöli az almás sütit.
many child hates the apple pie
'Many children hate apple pie.'

B: Tévedsz, sok gyerek ˋszereti az almás sütit.
'You're wrong, many children like apple pie.'

In a context of this type, the constituent that appears in topic position is typically characterized by the same rising intonation contour as contrastive topics, even if the sentence does not have the contrast implicature previously related to contrastive topics.

A type of topic with an unambiguously contrastive topic intonation and a special interpretation is the quantificational NP with a narrow scope over its associate.[13] According to É. Kiss and Gyuris (2003), the main function of their rising intonation contour is to individuate a property denoted by a quantificational NP.

As (39) illustrates, such a quantificational NP can function as an answer to a question about quantities:

(39) Q: Kit csókolt meg legalább kettő lány?
who-acc kissed pfx at least two girl
'Who was kissed by at least two girls?'

A: ˇLegalább kettő lány ˋPétert csókolta meg.
at least two girl Peter-acc kissd pfx
'It is Peter who was kissed by at least two girls.'

The examples above show that the denomination 'contrastive topic' covers an extremely heterogeneous class of data in Hungarian: there are many cases that cannot be differentiated from the ordinary topic, and the cluster of prosodic and semantic properties usually related to con-

[13] Another type of expression that can never appear in topic position without the characteristic intonation contour of the contrastive topic is the bare common noun, functioning normally as a verbal modifier in Hungarian. According to É. Kiss (2000), the contrastive intonation and interpretation of such NPs is a way of individuating the property they denote.

trastive topics does not hold for all cases. In section 4, we will examine how to characterize this heterogeneous category with respect to topicality.

3. Categorical and thetic judgments: the origins of the distinction

I assume that most of the problems discussed in section 2 can be traced back to a confusion about the notions of topic and logical subject on the one hand, and the nature of the categorical–thetic distinction, on the other. In fact, the current approaches to the topic in Hungarian theoretical linguistics take this notion as synonymous with that logical subject.

It is a well-known fact that the notion of topic can be defined in several ways. Most of the relevant definitions are based on the notion of "aboutness" and/or that of "givenness". The topic definition currently used in order to explain sentence structure in Hungarian is essentially based on "aboutness"; that is, a constituent occupying a certain syntactic position will be interpreted as the topic of the sentence independently of the context in which the sentence is uttered. On the other hand, the distinction between thetic and categorical sentences, as is demonstrated by the question-tests used to determine whether a sentence is thetic or categorical, is strongly related to the contextual properties of sentences.

The interpretation of the terms "categorical" and "thetic" has altered considerably since their introduction by Brentano and Marty. According to the theory of (Brentano 1973 [1874]) and Marty (1918), a categorical judgment is a double cognitive act, which consists of the recognition of a (logical) subject and the affirmation or denial of what is expressed by the predicate about the subject:

(40) Diese Blume ist blau.
'This flower is blue.'

(41) Mein Bruder ist abgereist.
'My brother has left.'

A thetic judgment is a logically simple judgment consisting of the act of recognizing or rejecting the content of a judgment:

(42) Es regnet.
'It's raining.'

(43) Gott ist.
'God exists.'

As (40)–(43) show—and this is also made clear by the authors—there is no direct relationship between the grammatical structure of a sentence and the type of judgment it expresses. Notice furthermore that Brentano and Marty have analysed decontextualized sentences only. They have claimed that the type of judgment expressed by a sentence does not depend on the context but on the logical structure of the sentence;[14] consequently, the question of a possible ambiguity between the thetic and the categorical reading (in the case of (41), for instance) does not even arise in their theory. Moreover, as Kuroda (1972) notes, Brentano and Marty do not consider sentences like (44) to express a thetic judgment:

(44) Ein Hund rennt.
'A dog runs.'

The great revival of the notions of categorical and thetic judgment and their application to the analysis of particular languages is in fact due to Kuroda (1972) as well as to Kuno (1972). Kuroda (1972) re-evaluates Brentano and Marty's logical theory on the basis of linguistic considerations and he points out that certain morphological characteristics of Japanese can be explained by means of the distinction between the two types of judgment. Sentence (45a)—in which the particle *ga* is attached to the noun *inu* 'dog'—corresponds to a thetic judgment, and sentence (45b)—in which the particle *wa* follows the noun *inu*—expresses a categorical one:

(45) (a) Inu ga hasitte iru.
'A/the dog is running.'
(b) Inu wa hasitte iru.
'The dog is running.'

Although Kuroda does not consider the morpheme *wa* as a topic marker (he keeps the original term 'logical subject' in order to avoid the confusion created by the various approaches of the notion of topic), he characterizes

[14] Notice also that, as Kuroda (1972) points out, the Japanese version of the sentence *Gott ist* 'God exists' includes the topic marker *wa* (*kami wa sonzai suru*); hence it would be considered a categorical sentence in Japanese. The same is true for the Hungarian version, *Isten létezik.*

the difference between (45a) and (b) in terms of contextual properties. He points out that sentence (45a) is used in a context where no dog was mentioned before, while sentence (45b) is uttered if the identity of the dog is already established in the preceding context. Furthermore, in the case of a thetic reading, both the definite and the indefinite interpretations are available for *inu*; for the categorical reading, however, only the definite interpretation is possible. This also shows that there is an important restriction in Japanese with respect to the referential properties of a *wa*-marked expression: it must be a definite NP or, as Kuno (1972) points out, if an NP other that a definite one (a quantificational NP, for instance) is *wa*-marked, it must be uttered with a special intonation and must be interpreted as implying a contrast.

In Lambrecht (1994), the thetic–categorical distinction is clearly analysed on pragmatic rather than logical grounds. In his approach, the topic of the sentence is always identified with respect to a particular discourse, that is, contextual determinacy plays a crucial role in the notion of topic he adopts. As for the possible interpretations of the thetic–categorical distinction, he argues in favor of an information structure approach to the thetic–categorical contrast which is based on pragmatic and not logical categories. He considers a sentence with topic–comment articulation to represent a categorical judgment and a sentence without such an articulation (i.e., a topicless sentence) to represent a thetic judgment. On the other hand, Lambrecht points out that the same syntactic structure, expressing the same logical proposition, can have different information structures in different discourse contexts. For example, sentence (46a) is compatible with both of the questions (46b) and (c):

(46) (a) The children went to school. (Lambrecht 1994, 121)
(b) What did the children do next?
(c) What happened?

Nothing in the syntactic or semantic structure of (46a) determines its information structure. The categorical or thetic character of the sentence is clearly a matter of pragmatics and must be dissociated from its syntactic and logical properties.

The approaches presented in this section show that the thetic–categorical contrast, which was originally established on logical grounds, has in recent decades become a distinction that concerns the pragmatic rather than the logical structure of the sentence. The problem with the notion of topic adopted in Hungarian theoretical linguistics is that it is defined

by means of the notion of logical subject elaborated in the original (logical) theory of Brentano and Marty and not on the basis of the current (pragmatic) interpretation of the thetic–categorical distinction.

4. The proposal

The problems discussed in section **2** can be solved if we make a clear distinction between the notions of topic and logical subject on the basis of the assumption that these notions belong to two distinct levels of sentence structure. I will assume, following Kiefer (1977) and Gécseg–Kiefer (forthcoming), that a sentence can be analysed on (at least) three levels: on the grammatical, semantico-logical and pragmatic levels. On the first it is grammatical relations that are defined, such as the grammatical subject–grammatical predicate relationship. The second level contains context-free logical relations such as the logical subject–logical predicate relationship and the third level is the site of pragmatic relations that are defined with respect to the particular context in which the sentence is uttered. One of the main relationships defined on this level is that of topic and comment.

4.1. Logical subject and topic

According to the basic claim of this paper, word order in Hungarian is determined by semantico-logical rather than pragmatic articulation. In this approach the notion of logical subject is exclusively based on aboutness; that is, a constituent occupying the position of logical subject denotes a referent about which something is asserted by the logical predicate. This referent can even be completely unidentified both by the speaker and the listener of the utterance. On the other hand, the logical subject of the sentence can play the role of the topic as well (although this is not necessarily the case), with respect to a certain discourse, if the identifiability conditions are met. We can define the notions of logical subject and topic as follows:

(47) (a) Logical subject:
The logical subject of a sentence refers to an individual (or group of individuals) about which something is asserted by the logical predicate.

(b) Topic:
The topic of the sentence refers to an individual (or group of individuals) fully identified by the participants of the discourse and holding a relation of aboutness with the proposition expressed by the sentence with respect to a particular discourse.

As the definitions (47a–b) suggest, there is a certain relationship between the two notions; that is, both of them are expressed in terms of aboutness. If a constituent appearing in topic position denotes a topic referent in the sense of (47b), this constituent is at the same time interpreted as denoting the logical subject of the proposition expressed by the sentence. On the other hand, a constituent in topic position can play the role of the logical subject without denoting a topic in a particular context. Consider for instance (25), repeated here as (48):

(48) Mari beteg/ megbetegedett.
Mary ill became-ill-pfx
'Mary is ill/has got ill.'

Sentence (48) can be taken to assert something about Mary or something about a particular situation, depending on the context. However, as we saw in section **2.2**, the syntactic structure (i.e., the word order) of the sentence is the same with the two interpretations. We can solve this puzzle by assuming that in any case there is an assertion about Mary but in some contexts the sentence pragmatically asserts something about a particular situation by means of a proposition which asserts something about Mary. The divergence of the articulations on the two levels (i.e., logical and pragmatic) is reflected by the prosodic properties of the sentence as well: in case of a thetic (i.e., topicless) reading, there is no intonational boundary between the logical subject and the logical predicate.

On the other hand, if the topic position of a sentence is occupied by a non-specific indefinite expression, the latter cannot play the role of the topic since it denotes a completely unidentified referent and, as a consequence, does not meet any condition of topicality. However, the sentence structure reflects that something is being asserted here about a(n unidentified) referent. In this sense we can claim that such a topicless sentence can have a logical subject–logical predicate articulation. This is the case of (29), repeated here as (49):

(49) (a) Egy kutya megharapott egy járókelőt.
a dog bit a passer-by-acc
'A dog has bitten a passer-by.'

(b) Egy járókelőt megharapott egy kutya.
a passer-by-acc bit a dog
'A passer-by has been bitten by a dog.'

As we saw in section **2.3**, the most natural reading of the two sentences above is about a completely unknown dog and a completely unknown passer-by. Consequently, (49a, b) are thetic sentences, in the same way as (49c) and (49d) are thetic sentences:

(49) (c) Megharapott egy kutya egy járókelőt.
bit-pfx a dog a passer-by-acc

(d) Megharapott egy járókelőt egy kutya.
bit-pfx a passer-by-acc a dog
'A dog has bitten a passer-by.'

The situation described by (49a–d) is the same in each case. Nevertheless, we have two basic variants of word order, an NP-initial and a verb-initial order, where the latter corresponds to a complex logical predicate structure and the former to a logical subject–logical predicate structure. In other words, each of the four sentences asserts something about a particular event, this event is presented in (49a) as an assertion about a dog, in (49b) as an assertion about a passer-by and in (49c–d) as an event of biting of a passer-by by a dog.

4.2. Identifying the topic

Since the topic-tests commonly used in Hungarian syntax (i.e., the sentence adverbial test and the prosodic criterion)[15] do not make reference to the particular context in which the sentence is uttered, it is easy to demonstrate that these tests serve to identify the logical subject rather than the topic (except for the cases when these notions overlap).

[15] For a detailed analysis of the sentence adverbial test and the prosodic criterion, see Gécseg–Kiefer (forthcoming).

In the previous sections of this paper another classical topic test, namely the question test[16] was implicitly applied. This test consists of the identification of an appropriate context in which a particular constituent functions as the topic of the sentence. The principle underlying this method can be formulated as follows:

(50) The topic part of a sentence is a (phonetically realized or non-realized) element of the sentence that is coreferential with the logical subject[17] of a preceding *wh*-question.

Principle (50) is in fact a very strong condition on topicality, because it requires that the topic referent of the sentence should already be introduced in the discourse. In other words, the topic referent of the sentence is an individual that a possible preceding question asks something about.

It is interesting in this respect to see about what kind of entities one can raise a well-formed question, i.e., what kind of NPs can occur in a *wh*-question. Consider the questions in (51):

(51) (a) Mit csinál Mari?
'What is Mary doing?'

(b)??Mit csinál az egyik gyerek?
'What is one of the children doing?'

(c)??Mit csinál három gyerek?
'What are three children doing?'

(d) *Mit csinál mindegyik gyerek?
'What is each child doing?'

(e) *Mit csinál sok gyerek?
'What are many children doing?'

(f) *Mit csinál valaki?
'What is somebody doing?'

[16] The question test can only be applied to non-negated sentences. On the other hand, Gécseg–Kiefer (forthcoming) makes use of the distinction between sentential negation and predicate negation in order to identify the topic constituent of the sentence.

[17] This principle is based on the hypothesis that a subclass of *wh*-questions has a binary structure, i.e., it can be divided into two parts: the first part is constituted by a *wh*-phrase and a verb and corresponds to what is asked, while the second part is constituted by the postverbal material and corresponds to what is being asked about. In this sense, we can say that such questions typically have a logical predicate–logical subject structure.

Among the six interrogative sentences above only (51a), a question about the referent of a definite NP (*Mari*) is totally well-formed. As for (51b) and (c), they are acceptable only in a situation where the speaker notices that one of the children (or three children) is/are doing something that the other children are not doing. But even in these cases one cannot answer them by means of sentences like (52a) and (b):

(52) (a) #Az egyik gyerek könyvet olvas.
'One of the children is reading a book.'
(b) #Három gyerek könyvet olvas.
'Three children are reading a book.'

Only answers of the type exemplified in (53), where the topic is a phonetically non-realized personal pronoun, are available for them:

(53) (a) Könyvet olvas.
'He is reading a book.'
(b) Könyvet olvasnak.
'They are reading a book.'

As for questions (51d–g), they are odd in any context. One can attempt to explain the ill-formedness of (51d) by the fact that *mindegyik gyerek* 'each child' denotes a universal quantifier that belongs to the predicate part of the sentence. However, this explanation does not hold for (51e–f): the postverbal constituent of these questions would obviously occupy the topic position in the corresponding assertive sentences. The problem with them is that sentences (51b–c) and (51e) are formulated about referents that have not been introduced to the discourse yet. Discourse-linkedness condition is not sufficient here: even if the NPs az *egyik gyerek* 'one of the children', *három gyerek* 'three children' and *sok gyerek* 'many children' can be taken to denote a member (or a subset) of an identified set of children, the referents of such indefinite NPs are not sufficiently identified to be able to serve as a subject for a subsequent predication.

As for (51f), the ill-formedness of this question comes from the fact that something is asked here about a completely unknown referent. This implies that if *valaki* 'somebody' has non-specific reading, the only question type that can precede a sentence of the form *valaki...* is the "thetic question" type like "What happened?", for instance.

Consider now question (54) and possible answers to it in (55):

(54) Mit csinálnak a gyerekek?
'What are the children doing?'

(55) (a) A gyerekek könyvet olvasnak./Könyvet olvasnak.
'The children are reading a book/They are reading a book'

(b) Az egyik gyerek könyvet olvas./(Az) egyikük könyvet olvas.
'One of the children is reading a book./One of them is reading a book.'

(c) Három gyerek könyvet olvas./Hárman könyvet olvasnak.
'Three children are reading a book./Three of them are reading a book.'

(d) Sok gyerek könyvet olvas./Sokan könyvet olvasnak.
'Many children are reading a book./Many of them are reading a book.'

(e) Mindegyik gyerek könyvet olvas./Mindegyikük könyvet olvas.
'Each child is reading a book./Each of them is reading a book.'

Question (54) is a perfect context for each sentence in (55). The members of the question–answer pairs are linked to each other by means of a contextually determined set of children, denoted by the NP *a gyerekek* 'the children' in (54) and (55a), by the common noun *gyerekek* 'children' contained in the sentence-initial NPs or by the implicit pronominal NPs in (55b–e). On the basis of the question-test we can conclude that in the sentences (55a–e) not only the predicate *könyvet olvas(nak)* 'is/are reading a book' brings new information about a referent referred to in the preceding question, but the determiner of the sentence-initial NP as well. In the context created by question (54), (55b–d) assert about an identified set of children that there is a particular member (or a subset) of this set for which the main predication of the sentence holds. In other words, these sentences contain a secondary predication of existence related to the indefinite character of the logical subject NP in them. As for (55e), we can maintain the generally accepted claim that this sentence has no logical subject and must be taken as a complex predicate. At the same time, its discourse-linked character shows that something is being asserted here about a set of children; that is, the nominal part of the universally quantified NP *mindegyik gyerek* 'each child' can be identified as the topic of the sentence. Since the topic does not correspond to a (phrasal) constituent here, we can conclude that this sentence has no topic–comment articulation on the grammatical level.

4.3. Applying the question-test to contrastive topic

We saw in section **4.2** that if we define topicality in terms of givenness and aboutness (in conformity with current assumptions about the topic appearing in categorical sentences), a strict version of question-test should be applied in order to identify the topic in a particular sentence uttered in a particular context. The application of this question-test led to the conclusion that among the definite and indefinite NPs appearing in topic position, only definite NPs can play the role of the topic. On the other hand, the topic expression of the sentence does not always correspond to a (phrasal) constituent in topic position, but can sometimes be identified as the nominal part of an indefinite (or even quantificational) NP, if the common noun contained in such NPs denotes a contextually determined set of individuals.

In this section, we examine whether the question-test defined for the "non-marked case", i.e., ordinary topic, can be applied for contrastive topic as well.

It was pointed out in section **2.4** that the term 'contrastive topic' refers to a class of linguistic objects characterized by a rather heterogeneous behavior. The classical cases, that of a definite NP functioning as a contrastive topic is relatively uncontroversial: the rising intonation contour of such NPs goes in pair with an implicature of contrast, which is normally absent in ordinary topics. On the other hand, the necessary partitive reading of specific indefinites and certain quantificational expressions makes it difficult to distinguish between "ordinary" and "contrastive" topic function of an expression occupying a topic position in a sentence. Finally, there is a subclass of NPs—namely quantificational NPs—that can never function as ordinary topics; they can nevertheless appear in topic position with the characteristic intonation contour of contrastive topics. At the same time, in topic position they lose their ordinary scope properties by having narrow scope over the operators they precede. As it was pointed out by É. Kiss–Gyuris (2003), the main function of the rising intonation contour of such quantificational NPs is not to imply a contrast, but to individuate the property denoted by the NP.

The first generalization we can make about sentences containing an NP with rising intonation in topic position is that they can never be thetic sentences, since they can never serve as answers to questions of the type "What happened?".

Let us examine then by means of the question-test whether the "contrastive topic" NP as a whole or only it's nominal part can be taken to be the topic of the sentence.

Consider again question (54), repeated here as (56), and the sentences in (57), serving as possible answers to (56):

(56) Mit csinálnak a gyerekek?
'What are the children doing?'

(57) (a) ˇA gyerekek/ˇŐk `könyvet olvasnak (, a ˇfelnőttek viszont `sétálni mennének).
'The children/They are reading a book (but the adults would rather go for a walk).'

(b) ˇMari `könyvet olvas (, ˇFeri pedig `tévét néz).
'Mary is reading a book (and Feri is watching TV).'

(c) ˇAz egyik gyerek `könyvet olvas./ˇ(Az) egyikük `könyvet olvas.
'One of the children is reading a book./One of them is reading a book.'

(d) ˇHárom gyerek `könyvet olvas./ˇHárman `könyvet olvasnak.
'Three children are reading a book./Three of them are reading a book.'

(e) ˇSok gyerek `könyvet olvas./ˇSokan `könyvet olvasnak.
'Many children are reading a book./Many of them are reading a book.'

(f) ˇA legtöbb gyerek `könyvet olvas./ˇA legtöbben `könyvet olvasnak.
'Most children are reading a book./Most of them are reading a book.'

The intonation marks indicate that the intended reading of the sentences in (57) is a contrastive topic reading. However, whereas the contrastive intonation and interpretation in (57a) clearly distinguishes this sentence from the corresponding (55a) in which the topic position is filled by an ordinary topic, it is very hard to make any prosodic or interpretational difference between (57c–e) and their ordinary topic counterparts (55b–d). On the other hand, the contextual connection between sentences (57c–e) and the preceding question (56) is established—just like in case of (55b–d)—by means of a relation of coreference between the NP *a gyerekek* 'the children' in (56) and the noun *gyerek* 'child' or an implicit pronominal-like element in (57c–e).

In sentence (57f) the contrastive topic is the quantificational NP *a legtöbb gyerek* 'most children'. As it was referred to in section **2.4**, this NP is sometimes considered as the logical subject (topic) of the sentence, and sometimes is taken to belong to the logical predicate. Furthermore, the fact that NPs of the type a *legtöbb N* 'most Ns' have partitive reading and

appear typically in sentences where they are associated to a constituent with prominent accent, it would be quite unnatural to try to make a difference between an ordinary and a contrastive topic function of such an NP. In any case, the "real" topic (in the sense of the definition (47b)) of the sentence (57f) should correspond to the noun *gyerek* contained in the NP *a legtöbb gyerek* 'most children', rather than to the NP as a whole.

As for (57b), the non-contrastive counterpart of this sentence would not be an appropriate answer to (56). Moreover, even though no formal link can be observed between the question and the answer, that is, there is no strict coreference between an expression in the question and an (explicit or implicit) element of the answer, the dialogue is coherent. Its coherence is in fact assured by the presupposition (shared both by the speaker and the listener of the utterance) that the referent of the contrastive topic constituent *Mari* is already identified as a member of a contextually determined set of children. Contrary to specific indefinite expressions like those in (57c–e), the referent of the definite NP *Mari* is identified independently of the main predication of the sentence. Since in this particular context the sentence presupposes, instead of asserting, that the referent of *Mari* in the answer belongs to the set denoted by *a gyerekek* 'the children' in the question, this referential autonomy of the definite NP makes possible for it to function as a topic in conformity with the definition of topic given in (47b).

One could be wondering what kind of question can be related to the narrow scope reading of quantificational expressions occurring in (contrastive) topic position. Such a context was exemplified in (39), repeated here as (58):

(58) Q: Kit csókolt meg legalább kettő lány?
who-acc kissed pfx at least two girl
'Who was kissed by at least two girls?'

A: ˇLegalább kettő lány ˋPétert csókolta meg.
at least two girl Peter-acc kissd pfx
'It is Peter who was kissed by at least two girls.'

In the question part of (58), the NP *legalább kettő lány* 'at least two girls' is interpreted not as a group consisting of two girls, but as a property of being a group of girls with the cardinality of two. If the first interpretation were correct, (59) would be an appropriate answer to the question in (58):

(59) #Pétert csókolták meg.
'It was Peter that they kissed'

On the other hand, the formal and referential identity of the NP *legalább kettő lány* 'at least two girls' in the question and in the answer indicates that the reference of this NP is contextually established, it can therefore be identified as the topic of the sentence.[18]

5. Topic vs. logical subject: cross-linguistic evidence

The main assumption of this paper is that Hungarian is not a topic-prominent language, but a logical subject-prominent one. This claim could be strongly supported if we could find languages that display a pragmatic rather than a logical articulation. Such a language would contain contextually unambiguous sentences with respect to information structure and would strongly constrain the referential properties of pre-verbal NPs.

Reference was made in section **3** to Japanese where a *wa*-marked expression must be a definite NP and other types of *wa*-marked NPs must be interpreted as implying a contrast.

Another piece of evidence comes from a spoken version of French, called colloquial French.[19] It has been observed that colloquial French has a tendency to avoid SVO order systematically in sentences with a non-pronominal subject (Cadiot 1988; Lambrecht 1994; De Cat 2002). Most sentences with a lexical subject — and sometimes even sentences with a pronominal subject—are topicalized structures using (left or right) dislocation of the constituent that plays the role of the topic. A sentence that displays a topic–comment articulation is either a dislocated structure (cf. (60)) or a sentence with a pronominal subject (cf. (61)):

[18] In fact, there is no need to apply the question-test in order to demonstrate that quantificational NPs with a narrow scope reading play the role of the topic of the sentence. Such NPs denote quantificational properties and quantificational properties are assumed to have a fixed reference. They can be compared in this respect to generic NPs: their referent can be identified without being introduced to a previous context.

[19] For a detailed analysis of information structure in colloquial French, see De Cat (2002) and Gécseg (forthcoming).

(60) Ma soeur, elle est malade.
my sister she is ill
'My sister is ill.'

(61) Elle est malade.
'She is ill.'

Sentences (60) and (61) are categorical sentences in the sense that they serve as answers to questions about the referent of the dislocated NP in (60) and the pronominal subject in (61). (Several constituents can be topicalized in the same sentence; that is, a sentence can contain more than one topic if required by the context.)

As for thetic judgments, they are typically expressed in colloquial French by means of existential constructions with an expletive subject *il* (which is often elided in informal speech) or presentative constructions introduced by the morpheme *voilà/voici*:

(62) (Il) y a ma soeur qui est malade.
it-expl there has my sister who is ill
'My sister is ill.'

(63) Voilà/ (Il) y a Pierre qui arrive.
voilà/ it-expl there has Peter who arrives
'Peter is arriving.'

Sentences (62) and (63) can never be uttered as answers to questions about the referent of *ma soeur* 'my sister' or that of *Pierre*. They usually appear at the beginning of a conversation or in the context of questions of the type "What happened?" or "What is happening?".

On the other hand, the nature of topicalized constituents is strictly constrained in colloquial French: topicalization can only affect definite or generic NPs and even specific indefinites cannot be dislocated:

(64) (a) Ce roman, je l'ai lu avec plaisir.
'This novel I liked very much.'
(b) *Un de ces romans, je l'ai lu avec plaisir.
'One of these novels I liked very much.'

At the same time, it is possible to topicalize the nominal part of an indefinite or quantificational expression:

(65) (a) *Beaucoup de/*Plusieurs/*Deux/*Quelques romans, je les ai lus avec plaisir.
'[There are] a lot of/several/two/some novels I liked very much.'

(b) Ces romans, j'en ai lu beaucoup/plusieurs/deux/quelques uns.
'As for these novels, I've read many/several/two/some of them.'

(66) (a) *Tous les romans, je les ai lus avec plaisir.
'As for all of the novels, I liked [them] very much'

(b) Ces romans, je les ai tous lus avec plaisir.
'As for these novels, I liked all of them very much.'

In (65b) and (66b), topicalization affects only the nominal part of the quantified expressions, represented as a definite NP in topic position and the determiner remains *in situ.* These sentences can only be interpreted as asserting about a contextually determined set of novels that a certain quantity of them was appreciated by the speaker.

In certain conditions, not only can a definite NP be topicalized, but also an indefinite expression:

(67) Des chemises, j'en ai des `propres.
indef.art. shirts I of-them have indef.art. clean-pl.
'As for shirts, I have some that are clean.'

(68) Du vin, j'en ai du `bon.
indef.art. wine I of-them have indef.art. good
'As for wine, I have some that is tasty.'

(69) Des fautes, il en fait à la `douzaine.
indef.art. mistakes he of-them makes to the dozen
'As for mistakes, he makes dozens of them.'

In sentences (67)–(69), the indefinite NP in topic position is interpreted as the name of a property, comparable in this respect to a bare common noun in topic position in Hungarian (exemplified by (10) in section 1). In some contexts a contrastive reading can be associated with such sentences, but this is not necessarily the case. Nevertheless, the behavior of the indefinite topic is similar to that of contrastive topics in Hungarian because in colloquial French the indefinite topic appears in sentences containing a constituent that is uttered with a prominent stress. In the case of a French sentence, this stressed constituent typically occurs at the end of the sentence. On the other hand, the property denoted by the topicalized constituent in sentences like (67)–(69) is always linked to the previous

discourse, that is, the referent of this constituent meets not only the condition of full identifiability, but also that of contextual boundness.

The examples above show that the preverbal field in colloquial French contains a position reserved for constituents that play the role of the sentence topic. Topicalized sentences only appear in a context where the referent of the topic constituent is already introduced and fully identified. Non-topicalized sentences contain either a pronominal subject or the existential construction *(il) y a* or the presentative morpheme *voilà*. In the case of a pronominal subject, this latter always functions as a topic:

(70) Il est malade.
'He is ill.'

Since a personal pronoun denotes an individual already mentioned in the discourse or present in the situation, the subject pronoun *il* 'he' is interpreted as the topic of sentence (70), which asserts that the referent of this pronoun is ill.

As for existential and presentative constructions, they can be taken as topicless because they appear typically in contexts where something is being asserted not about an individual, but about an event or a situation.

We can conclude from these data that colloquial French is a topic-prominent language in the sense that in this version of French there is a structural position reserved exclusively for a constituent that plays the role of the topic where topicality is defined in terms of full identification, contextual boundness and aboutness.

6. Summary

The main purpose of this paper was to examine the properties of the so-called topic constituent in Hungarian sentences. It was pointed out that if we make a strict distinction between the notion of topic, defined on a pragmatic level with respect to possible contexts, and the notion of logical subject, defined on a decontextualized logico-semantic level, Hungarian should be taken as a logical subject-prominent rather than a topic-prominent language. The existence of topic-prominent languages as opposed to logical subject-prominent languages was proved through data from colloquial French. As for the so-called contrastive topic in Hungarian, the results of the analyses show that only a subclass of contrastive topic expressions meets the topicality conditions established in

this paper on pragmatic grounds, and other types of contrastive topic expressions, namely those that can hardly be differentiated from ordinary topics, display the properties of logical subjects rather than topics.

References

Brentano, Franz 1973. Psychology from an empirical point of view. Routledge, London. Translated by Antos Rancurello, D. B. Terrell and Linda McAlister from Psychologische vom empirischen Standpunkt (1874/1924).

Cadiot, Pierre 1988. Le thème comme synecdoque. In: Langue Française 78 : 9–26.

De Cat, Cécile 2002. French dislocation. Doctoral dissertation, University of York.

É. Kiss, Katalin 1994. Sentence structure and word order. In: Katalin É. Kiss – Ferenc Kiefer (eds): The syntactic structure of Hungarian (Syntax and semantics 27), 1–90. Academic Press, New York.

É. Kiss, Katalin 1998. Mondattan [Syntax]. In: Katalin É. Kiss – Ferenc Kiefer – Péter Siptár: Új magyar nyelvtan [A new Hungarian grammar], 15–184. Osiris Kiadó, Budapest.

É. Kiss, Katalin 2000. A [+ referáló] és [+ specifikus] jegyek ellenőrzése a kontrasztív topik esetében [The checking of [+ referential] and [+ specific] features in the case of contrastive topic]. In: László Büky – Márta Maleczki (eds): A mai magyar nyelv leírásának újabb módszerei IV [Recent methods in the description of contemporary Hungarian 4], 85–96. JATEPress, Szeged.

É. Kiss, Katalin 2002. The syntax of Hungarian. Oxford University Press, Oxford.

É. Kiss, Katalin – Beáta Gyuris 2003. Apparent scope inversion under the rise fall contour. In: Acta Linguistica Hungarica 50 : 371–404.

Enç, Mürvet 1991. The semantics of specificity. In: Linguistic Inquiry 22 : 1–25.

Gécseg, Zsuzsanna 2001. A kontrasztív topik szintaxisáról és szemantikájáról [On the syntax and semantics of contrastive topic]. In: Magyar Nyelv 97 : 283–292; 423–431.

Gécseg, Zsuzsanna forthcoming. A magyar és a francia mondat tagolásának alapelvei / Les principes de l'articulation de la phrase assertive dans le hongrois et dans le français parlé. In: Ilona Kassai (ed.): A francia és a magyar mondat információs szerkezete / Structure informationnelle de la phrase française et de la phrase hongroise. Lharmattan & Tinta Kiadó, Budapest.

Gécseg, Zsuzsanna – Ferenc Kiefer forthcoming. A new look at information structure in Hungarian. In: Linguistics.

Gyuris, Beáta 2002. The semantics of contrastive topic in Hungarian. Doctoral dissertation, ELTE, Budapest.

Kálmán, László 2001. Magyar leíró nyelvtan. Mondattan I (Segédkönyvek a nyelvészet tanulmányozásához VI) [Hungarian descriptive grammar. Syntax 1 (Supplementary textbooks on linguistics 6)]. Tinta Kiadó, Budapest.

Kiefer, Ferenc 1977. Functional sentence perspective and presuppositions. In: Acta Linguistica Hungarica 27 : 83–109.

Kuno, Sosumo 1972. Functional sentence perspective: A case study from Japanese and English. In: Linguistic Inquiry 3 : 269–320.

Kuroda, Sige-Yuki 1972. The categoric and thetic judgment: Evidence from Japanese syntax. In: Foundations of Language 9 : 153–85.

Lambrecht, Knut 1994. Information structure and sentence form. Cambridge University Press, Cambridge.

Maleczki, Márta 2001. Indefinite arguments in Hungarian. In: István Kenesei (ed.): Argument structure in Hungarian, 157–99. Akadémiai Kiadó, Budapest.

Maleczki, Márta 2003. Information structure, argument structure and typological variation. In: Katarzyna Jaszczolt – Ken Turner (eds): Meaning through language contrast, 223–44. John Benjamins, Amsterdam & Philadelphia.

Marty, Anton 1918. Gesammelte Schriften. Vol. II., part 1. Abteilung. Max Niemeyer Verlag, Halle.

Molnár, Valéria 1998. Topic in focus: The syntax, phonology, semantics and pragmatics of the so-called "contrastive topic" in Hungarian and German. In: Acta Linguistica Hungarica 45 : 89–166.

Szabolcsi, Anna 1997. Strategies for scope taking. In: Anna Szabolcsi (ed.): Ways of scope taking (SLAP 65), 109–155. Kluwer, Dordrecht.

Varga, László 1987. Hozzászólás egy hangsúly-tanulmányhoz [Remarks on a prosody study]. In: Nyelvtudományi Közlemények 89 : 53–66.

Acta Linguistica Hungarica, Vol. 53 (2), pp. 175–226 (2006)
DOI: 10.1556/ALing.53.2006.2.4

A PHONETICALLY-BASED APPROACH TO THE PHONOLOGY OF [v] IN HUNGARIAN

ZOLTÁN KISS – ZSUZSANNA BÁRKÁNYI

Research Institute for Linguistics
Hungarian Academy of Sciences
Benczúr utca 33.
H–1068 Budapest
Hungary
{cash|bzsu}@nytud.hu

Abstract: We propose a unified, surface-based functionalist analysis of the phonology of Hungarian *v*, which is shown to fare better than past generative formalist/representational models. The model introduced can account for the two-fold patterning of *v* with respect to voicing assimilation without evoking exceptional means. Furthermore, it can also explain certain asymmetries as well as graduality displayed by *v*'s phonotactic distribution, namely, that some clusters are more frequent in the lexicon, whereas others are marginal. The analysis is grounded in the aerodynamics of *v*'s articulation (which involves inherently contradictory targets) as well as in the relative perceptibility of its contrast in various contexts. It is shown with the help of quantitative experiments that *v*'s phonological patterning is directly derivable from these phonetic factors.

Keywords: phonetics-based phonology, phonotactics, fricatives, voicing, [v]

1. Introduction

The seemingly odd phonological behaviour of [v] in Hungarian has always attracted a lot of attention in the phonological literature. Most attention has focused on its two-fold patterning in voicing assimilation, namely, that it patterns with obstruents in being targeted by the process, but it behaves like sonorants as it does not trigger voicing assimilation. Related to this dynamic aspect, its static phonotactic distribution has also been of interest, as it also displays asymmetrical properties. In this paper, we

1216–8076/$ 20.00

provide a functionalist account which can explain [v]'s static and dynamic phonology in a unified manner, based solely on its surface phonetics (in particular, its aerodynamic properties as well as its relative perceptibility in various phonetic contexts).

Having introduced the relevant data in Hungarian (§2) and in some other languages (§4), as well as the inadequacies of past (formalist/representational) approaches (§3), we present the phonetically-based analysis in section 5. First, the basic functionalist principles that the analysis uses are set forth, then we concentrate on the more specific aerodynamic and perceptual properties of [v], as well as those of the contexts in which it occurs. Based on these basic phonetic principles, we introduce the most important predictions of the analysis concerning [v]'s (i) realizations, (ii) behaviour in voicing assimilation, and (iii) phonotactic patterning. Last, we put forth the results of an acoustic experiment whose primary aim was to characterize the surface realizations of [v] and then check the validity of the proposed predictions, which are shown to be directly related to these realizations.[1]

2. The distribution of [v] and the "Voicing Requirement"

Let us begin with the basic facts concerning [v]'s distribution in Hungarian monomorphemic two-member consonant clusters ([v] does not occur monomorphemically in clusters with more than two consonants in this language). Table 1 below displays the logical possibilities of [v]'s clustering ability in three environments: (i) intervocalically, (ii) word-finally (before a pause), and (iii) word-initially (after a pause). Example words are also provided; in some cases the type frequency of the cluster is low (in other words, it occurs in but a handful of words), such clusters are marked with numbers which represent in how many words they actually

[1] In the first portion of the paper, we simply use the IPA symbol [v] to refer to what is usually and traditionally described as the "voiced labiodental fricative". In the second half of the paper, the exact phonetic identity (and variants) of this sound will be made more explicit. Sometimes, however, the orthographic form *v* is used to refer to this consonant in general, without going into details as to its factual realizations. In most cases, we use the orthographic forms of Hungarian words without providing IPA transcription. The IPA transcription of the letters the interpretation of which is non-obvious are as follows: *ty* = [c], *gy* = [ɟ], *sz* = [s], *s* = [ʃ], *zs* = [ʒ], *c* = [ʦ], *cs* = [ʧ], *dzs* = [ʤ], *ny* = [ɲ]; *a* = [ɒ], *á* = [aː], *e* = [ɛ], *é* = [eː]. An acute accent over vowel letters signals length.

occur. For some clusters, there are no monomorphemic examples, but since they featured in the experiment to be discussed below, they have also been included in the table; these non-monomorphemic clusters are shown bracketed.

Table 1

The distribution of [v] in monomorphemic words (CC clusters)
(based on Siptár–Törkenczy 2000, 98f; 106f; 129f)

		p	t	c	k	b	d	j	g	ts	tʃ/dʒ/r	s	ʃ	v	z	ʒ	m	n	ɲ	l	r	j	h
V_V	1. VvCV						bovden 1		afgán 2											bóvli	sevró 2	szovjet 1	
	2. VCvV	lopva	hatvan	kotyvaszt 1	ekvivalens	szubvenció	medve	fegyver 3	dugvány 3			köszvény 3	fösvény		özvegy 2		nyamvadt 4	szenved	ponyva	tolvaj	árva 4	csajvadék 1	
#_	3. VvC#						(hívd)															(hívj)	
	4. VCv#						kedv 3										hamv 1	-szenv 2	könyv 2	nyelv 2	szerv	ölyv 1	
_#	5. #vCV																			vlach 2	Wrangler 1		
	6. #CvV		tviszt 1		kvarc		Dvorzsák 1		gvárdián 2	cvekedli 2		szvetter 3	svéd										

The glosses of the words in the table are as follows: *afgán* 'Afghan', *árva* 'orphan', *bovden* 'V-shaped belt', *bóvli* 'trash', *cvekedli* 'pasta with cabbage', *csajvadék* 'vagabond', *dugvány* 'cutting', *Dvorzsák* proper name, *ekvivalens* 'equivalent', *ellenszenv* 'aversion', *fegyver* 'weapon', *fösvény* 'miser', *hívd* 'call.2sg.def.imp.', *hívj* 'call.2sg.indef. imp.', *kedv* 'mood', *kotyvaszt* 'concoct', *könyv* 'book', *köszvény* 'arthritis', *kvarc* 'quartz', *lopva* 'furtively', *medve* 'bear', *nyamvadt* 'lousy', *nyelv* 'language', *ölyv* 'hawk', *özvegy* 'widow', *ponyva* 'canvas', *sevró* 'kid(skin)', *svéd* 'Swedish', *szenved* 'suffer', *szerv* 'organ', *szovjet* 'Soviet', *szubvenció* 'subsidy', *szvetter* 'cardigan', *tolvaj* 'thief', *tviszt* 'twist', *vlach* 'Vlachian', *Wrangler* 'Wrangler jeans'.

It is of course the blank cells of this table that constitute the cases of most interest, that is, the clusters that are missing in the language, as well as those whose type frequency is low. Clearly, the distribution of a single [v] is not restricted intervocalically. The moment the position on either the left- or the right-hand side is occupied by a consonant, distributional restrictions occur, with more restrictions cropping up pre-consonantally, as displayed by the low frequency numbers. Similar observations can be made with respect to the word-final as well as the word-initial position: [v]'s distribution is limited in the context of an adjacent consonant. These distributional effects are summed up in Table 2:

Table 2

The effect of the immediate environment on the distribution of [v]

left env.	[v]	right env.	example	effect on [v]'s distr.
V	[v]	V	*kavics*	no restrictions
C	[v]	V	*medve*	few restrictions
V	[v]	C	*bóvli*	restricted
V	[v]	#	*sav*	no restrictions
C	[v]	#	*kedv*	restricted
#	[v]	V	*vas*	no restrictions
#	[v]	C	*[v]rangler*	restricted

Glosses: *kavics* 'pebble', *medve* 'bear', *bóvli* 'trash', *sav* 'acid', *kedv* 'mood', *vas* 'iron', *Wrangler* 'Wrangler jeans'

Based on Table 2, we can set up a hierarchy of environments, which illustrates how the distribution of [v] is curtailed in various contexts; this is shown in (1), where "$X < Y$" means that Y is an environment where [v]'s distribution is more restricted than in environment X.

(1) {V__V, #__V, V__#} < C__V < {V__C, C__#, #__C}

Table 1 also (partly) illustrates one of the most salient properties of the phonology of Hungarian obstruents, what we may refer to as the **Voicing Requirement**. According to this requirement, two obstruents standing next to each other may not differ in voicing, that is, either (i) both are voiceless, or (ii) both are voiced (hence the lack of *[vt] or *[vk] clusters). This requirement embraces the whole of Hungarian obstruent phonology; that is to say, it applies morpheme-internally as well as over morpheme and word boundaries.[2] We are going to use the term "Voicing Requirement" (henceforth, "VR") to therefore cover both the "static"

[2] Vago (1980, 143) has proven to be wrong when he states that in Hungarian, voice assimilation is optional. It is only when flanking a relatively long pause that two neighbouring obstruents may have different voicing (see Szigetvári 1998a, 223 and Siptár–Törkenczy 2000, 198, who state that "[voicing assimilation] is postlexical (it applies across any type of boundary as long as no pause intervenes) but obligatory and non-rate-dependent"). Unfortunately, Vago's assertion caused other analysts to work with data with empirical inadequacies and consequently reach erroneous conclusions, such as Lombardi (1995; 1999); see, for example, Törkenczy (2000) as well as Kenstowicz et al. (2003) on Lombardi's problematic claims. It must also be noted that Vago does not call voicing assimilation optional in the

and "dynamic" aspects of the phenomenon, what the traditional literature calls **voicing agreement** morpheme-internally and **voicing assimilation** over morpheme and word boundaries.

Word-initial CC clusters constitute a sub-case, where VR is stricter: in this position, obstruent clusters in Hungarian are always voiceless.[3] If we consider [v] a voiced fricative (as the traditional approach does), then this segment is the only regular exception to this generalization, as we do find voiceless/voiced obstruent plus [v] clusters in this position:

(2) Word-initial obstruent plus [v] clusters

tviszt 'twist', *kvarc* 'quartz', *szvit* ([sv]) 'suit', *svéd* ([ʃv]) 'Swedish', *Dvorzsák* (proper name), *gvárdián* 'guardian', etc.

In this position thus, [v] patterns with sonorant consonants, which are free to occur here (see (4)).

The "dynamic" aspect of VR works the following way: if two obstruent segments with different values for voicing come to stand next to each other, it is always the second segment that determines the voicing of the first, thus it acts as the trigger of the regressive voicing assimilation. Clusters flanking a boundary between (i) a stem + suffix and (ii) two words (compound words as well as phrases) are affected (so long as no pause obtrudes). The phenomenon is iterative, that is, it can apply to its own output. (3) illustrates this with a few examples:

(3) Regressive voicing assimilation in Hungarian

(a) voicedness assimilated:

[t][b] → [db]: e.g., *hát-ba* 'back-illat.'; *két#barát* 'two friends'
[ʃ][b] → [ʒb]: e.g., *has-ba* 'stomach-illat.'; *hús#bolt* 'meat shop'

(b) voicelessness assimilated:

[b][t] → [pt]: e.g., *láb-tól* 'foot-abl.'; *láb#torna* 'foot exercise'
[z][t] → [st]: e.g., *víz-tól* 'water-abl.'; *víz#torony* 'water tower'

(c) voicing assimilation is right-to-left iterative:

[sk][b] → [zgb]: e.g., *groteszk-ben* 'grotesque-iness.'
[gd][t] → [ktt]: e.g., *smaragd-tól* 'emerald-abl.'

---

detailed discussion of the rule (pp. 34f), only when he lists it among other rules in the Appendix.

[3] Except for, perhaps, one word *dzéta* 'dzeta' (ζ), if [dz] is considered a cluster, but its status is not uncontroversial; see, for instance Siptár–Törkenczy (2000, 87ff).

Crucially, VR does not restrict obstruent–sonorant/sonorant–obstruent clusters. In monomorphemic words, both voiced as well as voiceless obstruents can stand next to a sonorant. Also, a sonorant will not voice a preceding voiceless obstruent and a voiceless obstruent will not devoice a preceding sonorant.[4]

(4) Obstruent–sonorant sequences (monomorphemic)
plakát 'poster', *próba* 'rehearsal', *tréfa* 'joke', *knédli* 'dumpling', *klarinét* 'clarinet', *krém* 'cream', *gnóm* 'gnome', *grafika* 'graphics', *friss* 'fresh', *szmog* 'smog', etc.

(5) Lack of regressive voicing assimilation in obstruent–sonorant sequences
(a) Sonorants are not devoiced:
[m][t] →[mt] (*[m̥t]): *rém-től* 'monster-abl.'
[l][t] →[lt] (*[l̥t]): *hal-tól* 'fish-abl.'
(b) Sonorants do not voice:
[p][n] →[pn] (*[bn]): *kép-nél* 'picture-adess.'
[s][n] →[sn] (*[zn]): *rész-nél* 'part-adess.'

Turning to [v], we can state that it — apparently — behaves asymmetrically with respect to VR: it undergoes devoicing (6a),[5] but does not trigger voicing (6b):

(6) [v] and the Voicing Requirement
(a) [v][t] →[ft]: *sav-tól* 'acid-abl.'
[v][h] →[fh]: *sav-hoz* 'acid-all.'
(b) [t][v] →[tv] (*[dv]): *két vár* 'two castles'
[p][v] →[pv] (*[bv]): *szép vár* 'nice castle'

Thus pre-obstruent ("coda")/target [v] behaves as an obstruent, while post-obstruent (and prevocalic) ("onset")/trigger [v] patterns with sonorants.

[4] This is true of the standard dialect of Hungarian, or "Educated Colloquial Hungarian (ECH)" (see Siptár–Törkenczy 2000, 3), the dialect of Hungarian which this paper discusses. However, there are dialects (in Western Transdanubia) where anterior nasal sonorants trigger voice assimilation: *kismiska* [kiʒmiʃkɒ] 'trifle', *hát nem* [haːdnɛm] 'well not', etc. (see Fodor 2003, 339). Interestingly, and — as we will see (Table 4) — perhaps not accidentally, it is these dialects that display voice assimilation triggered by [v], too, as well as word-final devoicing.

[5] As we have seen above, this aspect of [v]'s behaviour is statically satisfied by the fact that there are no monomorphemic words with a [v] plus voiceless obstruent cluster.

A further generalization, the last one here, that we can draw from Table 1 is the following: in word-final position, [v] patterns with obstruents as it can cluster with sonorants as the second consonant. Sonorants do not normally occupy such a position.[6]

(7) Word-final C[v] clusters (complete list)

[mv]: *hamv* 'ash' (normally occurs suffixed as in *hamvai* 'his/her ashes')
[nv]: *ellenszenv* 'aversion', *rokonszenv* 'sympathy'
[ɲv]: *könyv* 'book', *enyv* 'glue'
[lv]: *elv* 'principle', *nyelv* 'language'
[rv]: *terv* 'plan', *szerv* 'organ', *érv* 'argument', *konzerv* 'tinned food', *ismérv* 'criterion', *keserv* 'sorrow', *mérv* 'extent', *orv* 'vile', *örv* 'guise', *sérv* 'hernia', *szarv* 'horn'
[jv]: *ölyv* 'hawk'
[dv]: *kedv* 'mood', *nedv* 'fluid', *üdv* 'salvation'

It is true though that there are not many examples of sonorant plus *voiced* fricative clusters word-finally anyway, some of the examples that the literature cites are obsolete as free forms (e.g., *nemz* 'beget', *tömzs* ([mʒ]) 'lode'). The most frequent examples involve [r] (e.g., *borz* 'badger', *törzs* ([rʒ]) 'trunk', etc.); this is true of [v], too: [v] is rare after sonorants other than [r] in word-final clusters.

Note also the three examples in the last row of (7): they involve [d] as the first member of the cluster. This cluster is also rare (and exceptional if one considers the sonority sequencing violation), but as Siptár and Törkenczy (2000, 80) put it, "the [sonority sequencing] violation is at least not unprecedented [(see, e.g., *edz* 'train$_V$', *pedz* 'begin to understand')] if /v/ is a fricative (an obstruent), whereas if it is a sonorant, [words like *kedv* 'mood'] would violate the otherwise exceptionless generalization that (on the surface) no final cluster can consist of a sequence of obstruent plus sonorant."

The generalizations regarding [v]'s behaviour that we detailed above can thus be summed up as follows: prevocalic [v] in a syllable onset behaves as a sonorant, while a [v] syllabified in a coda patterns and thus behaves as an obstruent. The two types of behaviour are manifest in [v]'s distribution in CC clusters as well as its patterning with respect to the Voicing Requirement:

[6] Two exceptions include the coronal liquid [l], which marginally occurs after [r] and [j] in recent borrowings and names: *görl* 'chorus girl', *fájl* 'file', *geil* 'yucky'. [j] does occur after sonorants word-finally; however—and crucially, as we will see later on—, its phonetic manifestation is a noisy fricative [j̝] (which often devoices to [j̊]): *férj* 'husband', *szomj* 'thirst'.

(8) Sonorant behaviour of [v]

(a) phonotactics: can stand with obstruents word-initially, like the other sonorants:
tviszt 'twist', *kvarc* 'quartz' ~ *tréfa* 'joke', *klarinét* 'clarinet'

(b) VR: "trigger" (post-obstruent/prevocalic) [v] does not satisfy VR (even though it is voiced), like the other sonorants:
hatvan 'sixty' *[dv] ~ *hátra* 'backwards' *[dr], *paplan* 'duvet' *[bl]

(9) Obstruent behaviour of [v]

(a) phonotactics: can stand after sonorants word-finally, like the other obstruents:
könyv 'book', *terv* 'plan' ~ *vonz* 'attract', *torz* 'distorted'

(b) VR: "target" (pre-obstruent) [v] satisfies VR, like the other obstruents:
[v][t] → [ft]: *sav-tól* 'acid-abl.' ~ [z][t] → [st]: *láz-tól* 'fever-abl.'

Here we must note that in the Western Transdanubian ("Nyugat-Dunántúl") dialect of Hungarian, "sonorant" [v] can trigger voicing assimilation; in this variety, other sonorants (most notably anterior nasals—see footnote 4) can also act as voicing triggers, and so [v] patterns with sonorants even here. Examples are listed in (10a) ("ECH" stands for Educated Colloquial Hungarian):

(10) [v]'s voicing behaviour in the Western Transdanubian dialect of Hungarian (WTH) (data from Imre 1971 and Juhász 2003)

(a) [v] voices a previous voiceless obstruent:
borotva 'razor': WTH [dv] (ECH [tv])
rakva 'put.part': WTH [gv] (ECH [kv])
vasvilla 'iron fork': WTH [ʒv] (ECH [ʃv])

(b) [v] devoices after a voiceless obstruent:
ötven 'fifty': WTH [tf] (ECH [tv])
csukva 'closed': WTH [kf] (ECH [kv])
húsvét 'Easter': WTH [ʃf] (ECH [ʃv])

In (10b), we can see another strategy in voiceless obstruent–[v] clusters, namely the progressive assimilation of voicelessness. This latter strategy is common in Vas county and the northern areas of Zala county. This type of assimilation is said not to apply over word boundaries (including compound words). According to Imre (1971, 263) and Fodor (2003, 339), both strategies are noticeable in three areas of Western Transdanubia: Zala county, the southern regions of Hetés and in Baranya county, even in the speech of the same speaker.

Works describing the "Janus-faced" behaviour of [v] often draw a parallel between its two-fold patterning and its phonetic manifestation. For example, Siptár notes that "[[v]'s] phonetic realizations form a **continuous range** of smooth transitions from a thoroughbred **fricative** with a lot of noise of friction to weak and noiseless **approximant**" (1996, 83; emphasis ours). Also, as Siptár (*op.cit.*, 88) as well as Siptár–Törkenczy (2000, 80, footnote 7) state, "phonetically, the degree of friction seems to correspond nicely to the [two-fold] pattern [displayed by [v]]." Specifically, when [v] is realized very noisy, its behaviour patterns with that of obstruents; when it is realized as an approximant, it displays a sonorant-like behaviour. Both works list examples that are meant to illustrate this correspondence between behaviour and phonetic realization. Table 3 below is an attempt at interpreting these claims concerning [v]'s phonetic realizations.

Table 3

The allophonic realizations of [v]
(after Siptár 1996, 83 and Siptár–Törkenczy 2000, 80)[7]

most fricatival:	1.	Dv#	*terv* 'plan'	"obstruent"
		vT	*hívsz* 'you call', *óvtam* 'I protected'	
	2.	VvDV	*révbe* 'to port', *bóvli* 'junk'	
		Vv#	*sav* 'acid'	
	3.	VDvV	*medve* 'bear', *olvas* 'read'	"sonorant"
		VvV	*kova* 'flint'	
		#vV	*vér* 'blood'	
least fricatival:	4.	#CvV	*kvarc* 'quartz'	
		VTvV	*pitvar* 'porch'	

According to the table above, the allophones of [v] can be grouped into four sets with respect to their fricativalness: the members of group 1 display the most friction, those of 4 the least. Group 2 and 3 occupy an intermediate place between these two extremes. The double line separates the [v]'s according to their phonological behaviour: groups 1–2 behave like obstruents, whereas groups 3–4 as sonorants. Notice that the [v] in *terv* 'plan' is claimed by Siptár (1996) and Siptár–Törkenczy (2000) to have the same degree of friction as the assimilated (fully devoiced) [v] (usually alleged to be a [f]) of *hívsz* 'you call' and *óvtam* 'I protected'.

[7] D = voiced consonant (including sonorants), T = voiceless obstruent. "#" is meant to signal absolute word-final/-initial position (a pause).

It is not made explicit at all whether the [v] in *terv* (and *kedv, könyv, szenv* etc., see (7)) is also devoiced or not. Actually, what one can infer from the works cited is that it preserves its voicing—and also maintains its friction, because Hungarian is claimed to display **no word-final obstruent devoicing.**[8] This aspect of [v]'s phonetics (voicing vs. friction) will be crucial in the phonetically-based analysis to be presented here. A major problem of this classification is that it is only based on the intuition of the authors, it is not backed up by any experimental evidence whatsoever. The bulk of the paper will thus be devoted to testing these claims regarding [v]'s realization. It is, however, not difficult to see that there does seem to be a strong link between [v]'s phonological behaviour in Hungarian and its phonetics; the problem is that no theory has been able to make this link explicit and explanatory. The approach of this paper will try to prove that the link is valid and can be used to explain the seemingly odd behavior of [v].

At this point it is appropriate to mention two apparent discrepancies between [v]'s patterning and its phonetic realizations. As both Siptár (1996, 88) and Siptár–Törkenczy (2000, 203) mention, in post-obstruent position, when the obstruent in question is a labial stop ([p, b]),[9] [v] is said to be a rather strong/noisy fricative, yet it behaves phonologically as a sonorant because it does not trigger voicing assimilation on the previous [p]:

(11) V[pv]V and V[bv]V clusters
[pv]: *lopva* 'stealthily', *alapvető* 'basic', *képviselő* 'representative', etc. (*[bv])
[bv]: *dobva* 'throwing', *szabvány* 'standard', *szubvenció* 'subsidy', etc.

The other minor hitch concerns word-initial #[v]C clusters. Examples for such clusters are not easy to come by; nevertheless, they include *Wrangler* 'a pair of Wrangler jeans', *vlach* 'Vlachian' (also occurs as *Vlach*, personal name) and *Vladimir* (personal name). As both Siptár (1996) and Siptár–Törkenczy (2000) claim, the [v] in these words is a fricative phonetically, yet it again behaves as a sonorant since it does not voice a previous voiceless obstruent:

(12) *két Wrangler* 'two pairs of Wrangler jeans': [tvr] (*[dvr])
két vlach 'two Vlachians': [tvl] (*[dvl])

[8] This is true of ECH; see footnote 4, however.

[9] Petrova–Szentgyörgyi (2004) actually extend the environment to include all labials, thus also including [m].

Interestingly, this behaviour is mirrored in similar clusters, as in *#hr-* [xr]: *Hradzsin* 'Hradzin (castle in Prague)', *Hruscsov* 'Khruschev' are articulated with a fricatival [x], even though in non-cluster onsets *h* is realized as the glottal glide [h] (single [x] occurs in coda position: *doh* [dox] 'musty smell', *jacht* [jɒxt] 'jacht').[10] This fact is mentioned in Siptár (1996) and Siptár–Törkenczy (2000), but no formal link is made between the phonetic fact and the phonology (allophony) of these segments despite the apparent parallel.

To sum up, the most important questions that this paper seeks to answer are the following:

(13) (a) What are the factors that direct [v]'s **phonotactic patterning**? (see Tables 1 and 2, (8a), (9a))

(b) What are the factors that lie behind [v]'s behaviour with respect to the **Voicing Requirement**? (see (8b) and (9b))

(c) Are [v]'s phonotactics and its behaviour with respect to the Voicing Requirement **related**? If they are, what is the reason for that?

(d) What **acoustic phonetic features** characterize the various realizations of [v]?

(e) Are [v]'s phonotactics and its behaviour with respect to the Voicing Requirement related to its **phonetic characteristics**? If they are, in what ways and how can it be modelled?

3. Previous approaches to the phonology of Hungarian [v]

The apparently asymmetrical behaviour of Hungarian [v] has attracted a lot of attention in the phonological literature over the past decades. In this section, we will provide a brief overview of these past approaches. The common denominator in them is the use of the generative **formalist/representational** (thus non-functionalist) model. (14) enumerates the most common suggestions that have emerged:

(14) (a) positing two underlying segments (obstruent [v] as well as sonorant [v])

(b) sonorant [v] is underlying, obstruent [v] is derived (Szépe 1968; Vago 1980; Olsson 1992; Siptár 1994)

(c) obstruent [v] is underlying, sonorant [v] is derived

(d) underlying [v] has an intermediate sonority value (Barkaï–Horvath 1978)

[10] A detailed description (and an OT analysis) on the allophony of Hungarian *h* can be found in Siptár–Szentgyörgyi (2002) and Szentgyörgyi–Siptár (2005).

(e) the use of an additional binary feature [±transient] (Zsigri 1994)

(f) feature geometrical approach (Kornai 1994)

(g) the use of underspecification/syllable constituents (Siptár 1996; Siptár–Törkenczy 2000)

(h) non-derivational approaches: Government Phonology (Szigetvári 1998a;b; Cyran 1997; Cyran–Nilsson 1998); Head-Driven Phonology (Ritter 2000); Optimality Theory (Petrova–Szentgyörgyi 2004; Blaho 2005)

In this paper, we cannot offer a comprehensive review of all these proposals due to space restrictions; here we only focus on the most important aspects of them.[11]

The claim that the double-faced behaviour of [v] can be explained by assuming two distinct underlying segments is usually dismissed in the classical phonemic and generative literature very quickly on the following grounds. First, the distribution of either segment is predictable: they are in complementary distribution. Also, the two allophones ("obstruent [v]" and "sonorant [v]") are phonetically similar. This then constitutes a classic case for allophony, with **one** allophone to be abstracted away as the underlying phoneme of the allophonic alternation. Furthermore, the alternation is absolutely automatic: a word-final [v] (claimed to be phonetically a voiced fricative [v]) is realized as an approximant sonorant (often transcribed as [ʋ]) when, for example, a vowel follows it: *sav* [ʃɒv] ~ *savas* [ʃɒʋɒʃ] 'acid' ~ 'acidy'.

If we follow the reasoning of the previous paragraph, then one choice available is to assume [v] to be a **sonorant** underlyingly. One of the earliest classical generative accounts—and without doubt the most influential one—couched in these terms is Vago's (1980). In his binary distinctive feature system, this consonant is thus [+son]. The rewrite rule that Vago (*op.cit.*, 35) postulates to account for voicing assimilation is given in (15):

(15) Voicing Assimilation *à la* Vago (1980) (1)

$$[-\text{son}] \rightarrow [\alpha\,\text{voi}] \,/\, \underline{\quad} \; (\#) \begin{bmatrix} -\text{son} \\ \alpha\,\text{voi} \end{bmatrix}$$

By assuming [v] to be [+son], underlying [v]'s will thus not be affected by Voice Assimilation. This way Vago is successful at accounting for those [v]'s that behave like sonorants (which do not voice a previous obstruent). The problem, of course, concerns the ones that behave like obstruents:

[11] A more thorough review can be found in Kiss (in preparation).

those that devoice before a voiceless obstruent. To achieve this, he needs to introduce an additional voicing assimilation rule which only refers to [v]. The two rules can be collapsed, as shown in (16):

(16) Voicing Assimilation *à la* Vago (1980) (2)

$$\left\{ \begin{bmatrix} -\text{son} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ -\text{cor} \\ +\text{cont} \end{bmatrix} \right\} \rightarrow [\alpha \text{ voi}] \Big/ \underline{\quad} \; (\#) \begin{bmatrix} -\text{son} \\ \alpha \text{ voi} \end{bmatrix}$$

(Obstruents and [v] are assimilated to a following obstruent in voicing.)

Furthermore, Vago (*op.cit.*, 46, note 6) assumes a "low-level adjustment rule" that makes [f] derived from [v] by rule (16) switch from [+son] to [−son], a kind of redundancy rule according to which, all voiceless segments are automatically [−son], too.[12]

The stance taken by Vago (1980) embraces many subsequent approaches. Therefore, the inadequacies that Vago's analysis suffers from are also transferred to those models. In addition to the problematic aspects of the theoretical foundations (namely, those of the derivational, rewrite rule-based system), the analyses are also handicapped from being unable to provide a unified explanatory analysis of [v]. The formalist models of the past are usually capable of accounting for only **one** aspect of the two-faced patterning of this consonant. If, for instance, [v] is assumed to be a sonorant underlyingly (as is the case in Vago 1980), only its sonorant face is explained; so that its obstruent behaviour may be accounted for, arbitrary, stipulative and often exceptional measures are taken. They come in the shape of late "adjustment rules", radical feature changes, exceptional feature geometrical configurations, absolute neutralization rules, redundancy rules, which are posited to operate at the "systematic phonetic level" in the "phonetic interpretation module".[13] Heavy

[12] By assuming [v] to be a sonorant, Vago (1980) is also successful at capturing the static phonotactic aspect of [v]'s sonorant behaviour, viz., it can stand after obstruents in initial clusters (8a). However, a separate rule is still required to account for its static obstruent behaviour (9a). (Vago himself does not mention the static aspects of [v]'s phonology.)

[13] Olsson (1992), with his unorthodox "as if rules" (sonorant [v] behaves "as if" it were an obstruent [v] before a consonant and in word-final position), as well as Szépe (1968) using abstract [β] to underlie [v]'s surface allophones are perhaps the most typical examples of the unconstrainedness of generative derivational analyses.

use is made of this module (see especially Siptár's (1996) paper), but no principled analysis is given concerning this compartment of the grammar in any of the works that employ it.

Some works try to avoid the pitfalls of having to classify [v] either as a sonorant or an obstruent by assuming it to rather occupy an intermediate place. The earliest proposal along these lines is Barkaï–Horvath (1978). Their analysis draws on the idea of the **sonority hierarchy**; they propose the following sonority scale of segment classes and individual segments:

(17) Sonority hierarchy (Barkaï–Horvath 1978, 83)

stops	fricatives	[v]	nasals	[y][14]	[r]	[l]
1	2	3	4	5	6	7

Accordingly, the feature [sonorant] will have seven values, and the voicing assimilation rule (which they posit to be active in Hungarian, Hebrew and Russian) will refer to these values:

(18) Voicing Assimilation *à la* Barkaï–Horvath (1978)

$$[m\,\text{sonorant}] \rightarrow [\alpha\,\text{voice}] \,/\, \underline{\quad} \begin{bmatrix} n\,\text{sonorant} \\ \alpha\,\text{voice} \end{bmatrix}$$

where $m \leq 3$ and $n \leq 2$

That is, a segment with less than and including 3 on the sonority hierarchy scale undergoes voicing assimilation before any class which does not exceed 2 on this scale.

Certainly, this solution directly captures the intermediate attribute of [v] with respect to voicing assimilation ([v] undergoes the rule, but does not trigger it), but its phonetic reality is questionable (namely that [v] **always** has an absolute sonority value of "3", regardless of its position, context), and it is difficult to see how a multivalent feature system like this can be managed in a principled/constrained way (e.g., why is it only sonorancy that is multivalent?, how exactly does the phonological module operate on non-binary feature values?, etc.). Also, the way *SPE* (Chomky –Halle 1968) defines [sonorant] and the meaning of sonority does not seem to be fully compatible. And lastly, the formalism does not capture the phonological behaviour of [v] in an explanatory way.

[14] Barkaï and Horvath's [y] stands for IPA [j].

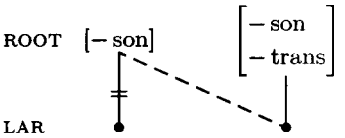
Another attempt to reflect on the intermediacy of [v] can be found in Zsigri (1994). He actually introduces a new feature [± transient], which divides the various segment classes, and [v], in the following way:

(19)

	obstruents	[v]	nasals	liquids/glides
[sonorant]	−	−	+	+
[transient]	−	+	−	+

This new feature is to be interpreted as follows: "[+ transient] is present in consonants during the articulation of which there is neither total closure nor fricatival noise" (Zsigri *op.cit.*, 19). Since this basically defines non-nasal sonorants, [nasal] is discarded in Zsigri's system. Zsigri's (1994) way of representing voice assimilation is shown in (20):

(20) Voicing assimilation *à la* Zsigri (1994, 69)



This way, voice assimilation can be represented by using a single rule and [v]'s behaviour with respect to it is captured correctly. It is, however, difficult to see how this new feature (and specifying [v] as [− son]) fares better than using the combination [±son, ± nasal], which divides the segments in (19) the same way, **if [v] is specified as [+ son]**. A step like that would of course take us back to square one, as it were, to the proposal of Vago (1980).[15] According to Zsigri (*op.cit.*, 22), [± transient] is also useful in expressing the phonotactic generalization concerning obstruent [v] as the second segment in "complex onsets". This is necessary in his framework as [v] is an obstruent, and so if it was rather classified as a sonorant, [± transient] would not be necessary, and with [± sonorant], the

[15] Zsigri (*op.cit.*, 67, 69) also claims that his system with the feature [± transient] and [− son] [v] is better at formulating the behaviour of [v] in Western Transdanubian voicing assimilation (see (10)): the rule in those dialects would only contain a triggering [− son], which will thus include [v]. The analyses where [v] is [+ son] would need to have three rules to express the dialectal assimilation. He also notes that positing different underlying [v]'s in the triggering vs. non-triggering dialects would be infelicitous, as there is no difference on the surface between the [v]'s. It is not clear why this is such a great problem, considering the capabilities of the derivational model (e.g., the use of late, "adjustment" rules), which are made great use of in his own analysis anyway.

relevant well-formedness condition could safely be expressed (as is done in Törkenczy 1994, see below).[16] Lastly, in addition to its dubious phonetic definition, the use of [± transient] cannot escape the difficulties of phonetic implementation: Zsigri, too, needs to resort to various default rules to ensure that a [f] devoiced from a [v] does not remain [+ transient] and a [v] voiced from a [f] becomes [+ transient].[17]

Even the frameworks that base themselves on more constrained representational foundations (such as autosegmental phonology, underspecification theory, the non-derivational frameworks: Government Phonology and Head-Driven Phonology (with their element theory and the giving up of the systematic phonetic level/module) as well as Optimality Theory) cannot fully cope with the behaviour of [v]. It actually turns out that in some cases (most notably in Government Phonology), the framework itself is too constrained to be able to offer a comprehensive account of the facts. As shown in Szigetvári (1998a;b), GP-based models face serious problems when explaining how an underlyingly sonorant [v] becomes an obstruent: if a strengthening process like this is assumed to involve the gaining of phonological elements, it is not clear at all how elements are acquired, from "nowhere", as it were. Furthermore, the fortition (obstruentization) of [v] is to take place in an environment (word-finally and pre-consonantally) where normally lenition phenomena occur.[18]

The problem of which variant of [v] is to be posited underlyingly apparently disappears in Optimality Theoretic analyses, owing to the OT architecture and its principle of "Richness of the Base", which states that there are no constraints on inputs; hence either [v], or [ʋ], or, an underspecified [v] can all be assumed underlyingly, and the grammar will still select the grammatical output. This stance is put forth in Petrova–

[16] [v]'s phonotactics in C__# (especially in *kedv*-like words, see (7)) remains to be a problem for the [± transient]-based analysis, but it has always been a problem for any formalist approach anyway.

[17] One more, independent, motivation for the use of [± transient] is its role in the vocalisation of transient consonants: transient sounds are hypothesized to be "capable of an easy alternation with vowels"; for details, see Zsigri (1994, 19).

[18] In Cyran–Nilsson (1998) [v]-obstruentization is represented as a change in the intrasegmental head-dependency relations ([w] {U} → [v] {U}), and so the problem of fortition not having a local source is circumvented. An alteration like that is in itself problematic (because of the change in the governing/licensing relations: a headless element is promoted for head status), but it still does not explain why it happens in a typical **lenition** site; furthermore, in their system, the voicing assimilation of [v] → [f] would still have to involve elements (such as "noise") appearing from nowhere.

Szentgyörgyi (2004), a paper focusing on [v] in both Hungarian and Russian, two languages where this consonant displays very similar behaviour. The paper starts off by setting up the constraints and their hierarchy, which can account for the voicing assimilation effects. If only these constraints were to be used for clusters with [v], then their system would generate ungrammatical outputs. This happens when [v] is prevocalic:[19] in the case of *ötven* [øtvɛn] 'fifty' the (wrong) output *[ødvɛn] is selected as optimal; the same happens in the case of *svéd* [ʃveːd] 'Swedish', where *[ʒveːd] is selected as optimal. To remedy this situation, they introduce a constraint which basically directs the conditions under which [v] can surface as a sonorant. This constraint, together with another one, which does not allow sonorants to voice, will now work well in selecting the actual grammatical candidates as optimal.[20] The two constraints referred to are shown below:

(21) (a) LAB SON
A voiced labiodental continuant is a sonorant if and only if it is followed by a sonorant.

(b) ID PRESYL SON VOICE
A segment before a syllabified sonorant in the output and its input correspondent must have identical specifications for voice.

A specialty of the constraint LAB SON in (21a) is that it does not refer to any prosodic constituent: it is only sensitive to the **linear**, immediate right-hand environment of [v], a rather different approach to other frameworks (among them Blaho's 2005). Additionally, notice that LAB SON particularizes that prevocalic [v] be a sonorant, whose behaviour is then directed by ID PRESYL SON VOICE (which will not let it voice a previous voiceless obstruent). This is clearly an attempt thus to connect [v]'s phonetic property with its phonological patterning with regard to the "Voicing Requirement".

The burdensome cases, the *lopva*-type and the *Wrangler*-type words, are also problematic for Petrova–Szentgyörgyi (2004). Their OT system will wrongly predict sonorant [ʊ] to surface in each of these cases. They acknowledge this issue, and propose the following constraint to resolve it.

[19] Note that Petrova–Szentgyörgyi (2004) posit inputs with a [−son] [v] in the following examples, thus, [øtvɛn] for *ötven* 'fifty' and [ʃveːd] for *svéd* 'Swedish'.

[20] We refer the reader to the paper of Petrova–Szentgyörgyi (2004) for the full evaluation analysis as the exact mechanism itself is not the main concern of the present discussion.

(22) *LABIAL + [v]
Labial consonants are not followed by a labial approximant.

Obviously, this (very specific) constraint must be ranked higher than LAB SON. Unfortunately, useful as it may seem, this analysis runs into problems — due to the constraint system already proposed to account for the other cases. Recall that LAB SON does not allow [v] to occur before a (syllabified) sonorant; therefore, neither [ʋ] nor [v] is allowed to surface in, say, *lopva*. Since Petrova and Szentgyörgyi's model also allows for underspecification, candidates with [v] underspecified for sonorancy (which they mark as [V]) are also generated, usefully enough, because in fact such an underspecified candidate is selected as optimal for *lopva* if the constraint in (22) is employed (and ranked high): *lo*[pV]*a*. So, a crucial role has to be played by the phonetic interpretation module—once again: this module will interpret an underspecified [V] as an obstruent by filling in the missing [– sonorant] feature.

To account for [v] + sonorant consonant clusters (in which a noisy [v] is supposed to surface),[21] Petrova and Szentgyörgyi make use of yet another constraint, *WC:[22]

(23) *WC
The feature [+ sonorant] may not be associated with a labiodental continuant before a (sonorant) consonant.

This highly ranked constraint (above *LABIAL + [v] and LAB SON) will not allow [ʋ] before sonorants, and so, again, a [V] will be selected as optimal ([v] + sonorant will be blocked by LAB SON), which phonetic interpretation will realize as a noisy, obstruent [v].

To sum up, Petrova and Szentgyörgyi's (2004) OT analysis of [v] in Hungarian, with its array of constraints as well as assumptions regarding input forms (underspecification), is (by and large) adequate descriptively. Phonetic factors do seem to be part of the analysis, most of the constraints

[21] The authors refer to **all** presonorant *v*'s here; but their claim that a *v* before a sonorant consonant is **always** realized as a noisy [v] is empirically wrong; see **§5.3.2**.

[22] *WC is a specialized version (relativized to a following sonorant) of the more general constraint *W, which bans voiced labial approximants (thus supposedly [ʋ], too). This latter constraint is posited by Petrova–Szentgyörgyi (2004) to account for the apparent fact that Russian *v* never surfaces as a sonorant [ʋ], but it still behaves in a dubious way phonologically. However, see Padgett (2002) and Lulich (2004), who show that Russian *v* does surface as a (narrow) approximant in some contexts.

are phonetically driven and claimed to be phonetically motivated, too.[23] However, phonetic facts are not the direct explanatory drives behind the phonological patterning in their model. It is therefore only apparently and partially phonetically-grounded: their crucial employment of underspecification emphatically shows that in some cases, a phonology-external compartment is supposed to arrive at a certain phonetic realization, which is phonologically relevant; a truly phonetically-based system would do the reverse: the phonetic realization of a segment (or rather, its phonetic characteristics and those of its linear surroundings) would be motivating phonological patterning **within the phonology**.[24] This would be especially fitting when there is a clear-cut linkage between behaviour and realization (like in the case of Hungarian [v]) in every position it finds itself. Notice also that Petrova and Szentgyörgyi's model crucially relies on syllable structure (see the underspecification of unsyllabified sonorants) and so the claim that in the case of **some** constraints, linearity (the avoidance of syllabic constituency) is an advantage cannot be accepted without reservations: a more compact (and truly phonetics-based) model would not, and perhaps should not, need to make reference to prosodic constituents, this would be **genuine** linearity. Also, [v]'s homogeneous behaviour is expressed as if it were the result of separate, independent constraints. In this paper, we will try to show that a single phonetically-based explanation can account for the static/dynamic behaviour of this consonant (and other, phonetically similar, ones) in the various positions it finds itself.

So far we have only been discussing past approaches to the analysis of [v]'s phonological behaviour in Hungarian voicing assimilation, but what about the generalizations concerning its phonotactics, those that we discussed at the beginning of the paper? As long as some of the phonotactic constraints concerning [v] can be related to what we have called the static manifestation of the "Voicing Requirement", then, of course, the previous analyses did tackle this aspect of [v]'s phonology, too. These

[23] Phonetic grounding of constraints in OT was a long-needed development in the history of the theory. See Kager (1999, 11f), who warns against positing constraints "lacking both typological motivation and phonetic grounding, even if there is compelling motivation for [them] from the language data under analysis." Indeed, phonetically-grounded phonology can be seen as a reaction against OT's highly suspicious arbitrary "parochial" and language-specific constraints. See also Wright (2004) on this issue.

[24] This is why Petrova and Szentgyörgyi's criticism of Padgett's (2002) relegating some phenomena of voicing assimilation in Russian to the phonetic component is not just.

analyses included [v]'s distribution in (i) intervocalic CC clusters and (ii) word-initial obstruent + [v] clusters. In the first case, the fact that no [v]–voiceless obstruent clusters exist in Hungarian is due to the Voicing Requirement as well as that [v] can only stand before voiced obstruents. Word-initial obstruent–[v] clusters have also been widely discussed, due to their seemingly exceptional behaviour regarding Voicing Requirement. But no analysis we have reviewed has stepped beyond these cases and attempted to provide a **unitary** account of the distributional facts. Moreover, none of the past approaches have offered an explanatory analysis on the **frequency** facts concerning [v]'s distribution, namely that some clusters are very frequent, others are rare or even downright impossible (this we may call **phonotactic graduality**). For instance, the models focusing on voicing assimilation did not account for the fact that although [v]–voiced obstruent clusters are the only expected clusters (and not [v]–voiceless clusters), the number of such sequences is nevertheless extremely low.

Two well-established works whose pronounced aim is to present a rather comprehensive account of the phonotactic facts of Hungarian are Törkenczy (1994) and Siptár–Törkenczy (2000). Unfortunately, they cannot escape the problems we have mentioned in the previous paragraph: [v]'s distributional facts as well as its behaviour concerning voicing assimilation are not treated in a unified manner; in fact, both works need to resort to exceptional devices.[25] Furthermore, their account on [v]'s phonotactics is only descriptively adequate, but does not offer an explanation; the nagging problem of the inability to account for the graduality of [v]'s well-formedness in consonant sequences still remains, namely that the type frequency of the relevant clusters is varied and cannot be captured by categorical means.

4. A crosslinguistic overview: [v] in Czech, Slovak and Russian

The double-faced behaviour of [v] is not unique to Hungarian. On the contrary, it is striking in how many languages (which may not be related genetically) [v] displays a similar, recurring phonological patterning. Here we provide a brief overview of three (related) other languages, where [v]'s

[25] They include simple (intuition-based) statements like "we regard syllable-initial [vl] [...] as exceptional" (Törkenczy *op.cit.*, 282f), as well as exceptional representational measures such as the introduction of "Sp(ecial) Licensing" (see Siptár–Törkenczy *op.cit.*, 140).

behaviour has been reported to be very similar to that in Hungarian: these languages are Czech, Slovak and Russian.[26]

(24) The Voicing Requirement in Czech, Slovak and Russian

(a) obstruent + obstruent clusters: regressive voicing assimilation

(i) Czech: [s][d] → [zd]: *s domem* 'with a house'
[z][p] → [sp]: *z pole* 'from a field'

(ii) Slovak: [s][b] → [zb]: *prosba* 'request'
[z][k] → [sk]: *z kina* 'from a cinema'

(iii) Russian: [t][b] → [db]: *ot-brosit^{j}* 'to throw aside'
[d][p] → [tp]: *pod-pisat^{j}* 'to sign'

(b) obstruent + sonorant clusters

(i) Czech: sonorants do not trigger voicing assimilation
[s][l] → [sl] (*[zl]): *s lesem* 'with a forest'
[s][m] → [sm] (*[zm]): *s mužem* 'with a man'

(ii) Slovak: sonorants can trigger voicing assimilation **across word boundaries**
[p][m] → [bm]: *chlap môže* 'a guy can'
cf. [p][m] → [pm] (*[bm]): *chlap-mi* 'guy.inst.pl'
cf. [k][n] → [kn] (*[gn]): *vlákno* 'fibre'

(iii) Russian: sonorants do not trigger voicing assimilation
[k][n^{j}] → [kn^{j}] (*[gn^{j}]): *kn^{j}az* 'prince'
[s^{j}][m] → [s^{j}m] (*[z^{j}m]): *pis^{j}ma* 'letter.pl'

(c) sonorant + obstruent clusters: no voicing assimilation in any of the three languages

(25) Word-final consonants and voicing in Czech, Slovak and Russian

(a) word-final obstruents are devoiced

(i) Czech: *mužem* [ʒ] 'man.inst.sg' ~ *muž* [ʃ] 'man.nom.sg'

(ii) Slovak: *zväzu* [z] 'union.gen.sg' ~ *zväz* [s] 'union.nom.sg'

(iii) Russian: *kniga* [g] 'book.nom.sg' ~ *knik* [k] 'book.gen.pl'

(b) word-final sonorants are not devoiced in any of the three languages[27]

[26] On the Czech and Slovak data, see Hall (2003; 2004). The Russian examples were taken from Padgett (2002). Note that in his examples, he suppresses predictable palatalization, vowel reduction, and other irrelevant surface variation. For more on the Russian data and the exact domain of the voicing phenomena, also consider Halle (1959); Hayes (1984); Petrova (2003); Petrova–Szentgyörgyi (2004). Languages other than the ones we focus on have also been said to show similar [v]-patterns, they include Hebrew (Barkaï–Horvath 1978), Ukrainian (Cyran–Nilsson 1998), Swedish, Romanian (Lombardi 1995), and some dialects of English, too (Petrova–Szentgyörgyi 2004).

[27] For Russian this claim is not uncontroversial. According to Padgett (2002), however, the process is at best only gradient and optional, just as much as the infa-

(26) The Voicing Requirement and [v] in Czech, Slovak and Russian

(a) [v] as potential target: regressive voicing assimilation (~obstruents)

(i) Czech: [v][p] → [fp]: *v pole* 'in a field'

(ii) Slovak: [v][t] → [ft]: *v tom* 'in that'

(iii) Russian: [v][k] → [fk]: *korovka* 'cow.dim'
(cf. *korovok* [v] 'cow.dim.gen.pl')

(b) [v] as potential trigger

(i) Czech: [v] does not trigger voicing assimilation (~sonorants)
[t][v] → [tv] (*[dv]): *tvůj* 'your'[28]

(ii) Slovak: [v] can trigger voicing assimilation **across word boundaries** (~sonorants)
[t][v] → [dv]: *brat vám (zavolá)* 'brother (calls you)'
cf. [t][v] → [tv] (*[dv]): *tvár* 'face'

(iii) Russian: [v] does not trigger voicing assimilation (~sonorants)
[t][v] → [tv] (*[dv]): *tver'* 'Tver'

(27) Word-final [v] in Czech, Slovak and Russian

(a) Czech: word-final [v] gets devoiced (~obstruents):
zpěvem [v] 'song.inst.sg' ~ *zpěv* [f] 'song.nom.sg'

(b) Slovak: word-final [v] is realized as [w]:
ústavom [v] 'institute.inst.pl' ~ *ústav* [w] 'institute.nom.sg'

(c) Russian: word-final [v] gets devoiced (~obstruents):
prava [v] 'right.fem' ~ *prav* [f] 'right.masc'

Clearly, [v] in Czech, Slovak and Russian displays the same behaviour as Hungarian [v] in consonant clusters with respect to voicing. There are three important differences we must point out, nonetheless. First, unlike these three languages, (Educated Colloquial) Hungarian has not been reported to show word-final obstruent devoicing (for [v] either). Second, Slovak appears to behave differently with respect to sonorant voicing: in this language, sonorants are said to voice voiceless obstruents before them across words and some morpheme boundaries. Lastly, Slovak is also special as in this language, word-final/pre-consonantal [v] is not realized as [f], but as [w], a type of lenition. The facts regarding voicing assimilation, word-final devoicing and the behaviour of [v] with respect to these two

mous transparency of Russian sonorants, as in *i*[s mts]*enska* 'out of Mcenk' vs. *i*[z o]*kna* 'out of the window'.

[28] In some dialects of Czech, [v] surfaces as a [f] after voiceless obstruents; this is the same behaviour as in the Western Transdanubian dialects of Hungarian (see (10)); further details can be found in Hall (2003; 2004).

phenomena are summarized in Table 4 for five languages (we also included the Western Transdanubian dialect of Hungarian (WTH) in the chart; "ECH" = Educated Colloquial Hungarian; "C" = Czech; "R" = Russian; "S" = Slovak). A "+" is used if the relevant phenomenon applies in the given language; "–" if it does not (but also consider the notes).

Table 4

Voicing effects and the behaviour of [v] in Hungarian, Czech, Slovak and Russian

	ECH	WTH	C	S	R
regr. voice ass. in obs.+obs. clusters	+	+	+	+	+
regr. voice ass. in obs.+son. clusters	–	(+)*	–	(+)†	–
final obs. devoicing	–	+	+	+	+
[v] as target (devoicing)	+	+	+	+	+
[v] as trigger (voicing)	–	(+)*	–	(+)†	–
[v] in final devoicing	–‡	+	+	(–)§	+

* Sources do not specify the exact domain of presonorant voicing.
† Only applies across specific morpheme boundaries.
‡ However, see the results of the experiment below.
§ [v] is not devoiced but lenited to [w].

The most important generalizations are thus as follows: (i) all languages display regressive voicing assimilation in obstruent–obstruent clusters; (ii) target [v] behaves in all languages as an obstruent; (iii) [v] only triggers voicing assimilation if the other sonorants do, too (and in the same morphological context); (iv) word-final [v] patterns with obstruents with respect to devoicing (with the exception of Slovak).[29] A last implication that we can draw from Table 4 is that if a language displays presonorant voicing (over word boundaries), then that language also has word-final devoicing (the reverse is not true, however). The connection between presonorant voicing and word-final devoicing seems to be an important factor in languages where sonorants do not contrast in voicing but they can be argued to demonstrate presonorant voicing.

[29] A more suitable generalization could be then to say that word-final [v] patterns with obstruents in that it is "unstable": it is prone to changing in this position. Actually, the analysis that we will be putting forth will predict what routes this change can take.

In the remainder of the paper, we will try to argue for an analysis that can account for these recurring generalizations regarding [v]. Even though we will focus on Hungarian, we will predict that the account can be extended to the languages tackled in this section, and other languages, too.[30]

5. A phonetically-based approach to [v]

5.1. Functionalism and phonetics in phonology

There is a growing body of evidence that **functional** factors previously thought to be external to grammar can nevertheless exert direct influence on it. These factors include such "low level effects" as speech **production** (articulation) and speech **perception**.[31] The basic idea that we pursue in this paper is that the phonetic (specifically, the aerodynamic and acoustic) properties of sounds can regulate their phonological patterning (including segmental distribution and allophony). The four most important functional principles that we make use of here are listed in (28):

(28) (a) creation of contrast;
(b) maximization of the number of contrasts;
(c) maximization of the perceptual distinctiveness of contrasts;
(d) minimization of articulatory effort.

The first of these principles is responsible for the creation of contrastive cognitive categories; by maximizing the number of contrastive categories (28b), the expressiveness of communication is enhanced by building up a substantial lexicon of categories. Principle (28c) accounts for the salience of the distinct basic categorical elements—according to it, categories must have acoustic properties that make them maximally different from each

[30] The works on [v] in the Slavic languages we have reviewed do not consider the **phonotactics** of [v], they only concentrate on its voicing behaviour (cf., however, Padgett's brief remarks on Russian obstruent–sonorant onset clusters (2002, 16)). We will, nevertheless, provide an account for Hungarian in which the phonotactic peculiarities can be explained as well. We anticipate that the analysis can also be transferred to the phonotactics of languages other than Hungarian.

[31] For a summary of functionalism, the role of articulation and speech perception in phonological theory, see, among others, Hume–Johnson (2001) and the references therein.

other perceptually; this principle thus requires segmental contrast to be sufficiently distinct and aims at **contrast preservation**. The last principle secures that the actual implementation (articulation) of the categories is to be carried out using as little energy as possible, a phonological consequence of which may be **contrast neutralization**. Notice that, as Flemming (2004) also shows, principle (28c) is inherently in conflict with both principles (b) and (d); thus, the phonological system is required to be such that it "weights" these principles and gives preference to one over the other.

The salience of a segmental contrast is **cue-based** and fundamentally depends on two factors: (i) the quality and quantity of the **inherent** acoustic cues of the given sound and (ii) the quality and quantity of the acoustic cues the sounds in its **immediate context** provide. Thus, the model makes no reference to segmental constituency (like the syllable), only the linear environment of a sound is considered. Segmental contrast is argued here to be licensed (upheld) in positions with sufficient and good-quality cues for that contrast to be perceived, recognized. Conversely, segmental contrast is predicted to be suspended, neutralized in badly-cued positions. This idea is expressed in the principle of **Licensing by Cue**:

(29) Licensing by Cue (Steriade 1999, 4):
The likelihood that distinctive values of the feature F will occur in a given context is a function of the relative perceptibility of the F-contrast in that context.

Another phonetic notion that we introduce is the **articulatory target**, which is defined as a given configuration that the articulatory system has to reach, realize. In view of Licensing by Cue, we argue that a given target in a badly-cued context is not realizable, and depending on the specific phonetic characteristics of the target and its context, particular **articulatory adjustments** are made, with the consequence of contrast neutralization. Neutralization in a position is to be understood in two ways: either only one set of phonetic properties is realized there (say, voicelessness and friction), which the cues can license, or the segment is lost altogether (deletion).

Steriade (1997) argues that for a given contrast (such as voicing–voicelessness of obstruents), the environments the contrast may occur in can be scaled hierarchically according to how well or how badly they cue the contrast in question. For the voicing contrast of obstruents, she sets up the following hierarchy of environments ("$X < Y$" means that Y is

an environment with less and worse perceptual cues than environment X (in other words, Y is more marked perceptually, less cue-able than X); "O" signals obstruents, "R" sonorants, and "#" a pause):

(30) Perceptual difficulty scale for the voicing contrast of obstruents:
V__R < {#__R, O__R} < V__# < V__O < {O__O, O__#, #__O}

According to this hierarchy, obstruent voicing is best perceived between a vowel and a sonorant. It is in this position that the relevant cues for the contrast are available in number and quality: closure voicing, closure duration, the duration of the vowel, F1 values in the vowel, burst duration and amplitude, VOT value, F0 and F1 values at the onset of voicing in the second sonorant. However, the cues to obstruent voicing are curtailed after a vowel and before an obstruent/pause, where important transition cues on the right are now missing. In the worst contexts, in the lack of a neighbouring sonorant, (O__O , O__#, #__O), the all-important transition cues are unavailable, merely the inherent cues are kept (the voicing and duration of closure).

We claim, along with Steriade (*ibid.*), that the hierarchy in (30) is "**inclusive**"; that is, if a contrast exists in a given position, then it also exists in those providing **better** cues for it. For instance, if a language has voicing contrast for obstruents in word-final position, it must also have that contrast after an obstruent and before a sonorant, word-initially, as well as between sonorants.[32]

5.2. The phonetics of voiced fricatives

We begin our phonetically-grounded, cue/target-based analysis of Hungarian [v] with considering the phonetics of voiced fricatives, in particular, their aerodynamic properties. For the articulatory system to target voicing and friction (turbulent noise) at the same time, an uneasy balance needs to be maintained.[33] First, high-amplitude turbulent noise requires a relatively high volume velocity of the air molecules as they pass through a channel. The faster the air moves, the louder the sound; thus, the ideal

[32] For more on the role of this monotonicity requirement in Hungarian and English phonotactics, see Rebrus–Trón (2002) and Kiss (2005).

[33] For the phonetic details of voicing and fricatives, see Stevens (1998, 477ff); Johnson (2003, 120ff); Jansen (2004, 40f) and Fuchs–Brunner (2005).

glottal configuration for a noisy fricative is when the glottis is **widely abducted** so that a high volume velocity airflow can pass through the vocal tract. Second, for a given rate of airflow, the narrower the channel, the louder the turbulent noise; according to Stevens (1998, 379ff), the glottal opening needs to be somewhat greater than the cross-sectional area of the supraglottal constriction. Lastly, turbulence is also produced when a jet of air hits an obstacle downstream. In the case of [s] and [ʃ], the upper and lower teeth constitute the obstacle, whereas for [f], it is the upper lip. If all these three factors are given, the sound produced is a high-intensity fricative.

On the other hand, for vocal fold vibration to take place, the glottis should be **closed** but loosely compressed, which is achieved by the air pressure to be **lower above the glottis** than below it.[34]

Based on the above, we can conclude that the production of high amplitude fricative noise and voicing involves contradictory articulatory gestures and aerodynamic events. Turbulent noise can be achieved by a widely abducted glottis and a constriction which is to be narrower than the glottal opening in the oral cavity. Both these factors result in an increase in intraoral pressure and a drop in subglottal pressure as more air builds up in the oral cavity at and behind the constriction area. An abducted glottis and a decrease in the transglottal pressure differential both remove the basic conditions for vocal fold vibration:

> "for the sake of continued voicing the oral pressure should be low, but for the sake of frication the oral pressure should be high. Meeting both of these requirements may be difficult. To the extent that the segment retains voicing it may be less of a fricative, and if it is a good fricative it runs the risk of

[34] As Jansen (2004, 36) estimates, the minimum transglottal pressure difference sufficient to maintain vocal cord vibration is around 200 Pa; to **initiate** fold vibration, this difference needs to be twice as much. A sound is defined to be **passively (modally) voiced** when during its articulation, a closed equilibrium position of the vocal folds and normal subglottal pressure is enough to initiate or maintain the physical conditions for vocal cord vibration (like in the case of sonorants). If passive voicing cannot be achieved (such as during the closure phase of stops), sounds are said to be **passively devoiced**. To overcome passive devoicing, a number of articulatory gestures (e.g., the enlarging of the oral tract volume) need to be implemented—this mechanism is referred to as **active voicing**, which characterizes ("truly") voiced obstruents. Conversely, **active devoicing** refers to situations in which sounds that would otherwise be passively voiced are nonetheless devoiced due to various additional articulatory means (e.g., the tensing of the vocal tract walls so that no passive expansion may occur and active decreasing of the size of the cavity behind the oral constriction by raising the larynx).

> being devoiced. In fact, the noise component of voiced fricatives is much less than that for voiceless fricatives [...] and on nonsibilant voiced fricatives ([β, v, ð, j, ɣ, ʁ]) is often so weak as to be barely detectable."
>
> (Ohala 1983, 201f)

In our analysis of Hungarian [v], we take this line of reasoning a step further. Our starting point is the aerodynamic incompatibility of the articulatory targets: our hypothesis is that the **maximal** implementation of **both** active voicing and strong frication at a labiodental place cannot be realized. The two targets can only be maintained if (i) voicing is passive/modal, due to a relatively open vocal tract, and (ii) frication is relatively low as a result of a wider constriction (much wider than that found in voiceless fricatives). The simultaneous realization of these targets can optimally be upheld in intervocalic, or more generally: intersonorant position, where passive voicing passes through the sequence of segments unimpeded (as in say [ava] or [avla]). Along these lines we claim that [v] is like a sonorant with respect to its voicing qualities but like a fricative in possessing turbulent noise, too, albeit at a lower level. This is in line with Padgett's (2002) analysis of Russian *v*, which he calls a **narrow approximant** and transcribes it as [ʋ̞] (from now on, we also adopt this symbol for notating Hungarian (prevocalic) *v*).[35] With respect to its noise qualities, this sound is claimed to stand between actively voiced and noisy [v] and passively voiced approximant [ʋ/w]. Between two sonorants, [ʋ̞] is predicted to display more formant structure and intensity and less turbulent noise than any other fricative (thus its noise is modulated by voicing), but less formant structure and intensity and more turbulent noise than any other ("wide") approximant.

Even though we base the starting point of our present analysis of [v] on aerodynamic grounds, a plausible perceptually-based account can also be suggested, as is done in Balise–Diehl (1994). They claim that the presence of voicing interferes with the perception of place cues in fricatives. Voicing-based laryngeal contrasts in fricative inventories tend to be neutralized because it is relatively hard to recover their place cues. They cite two pieces of evidence in support of this: the presence of voicing in a fricative reduces the amplitude of frication noise, which is an important cue for place contrast; furthermore, studies of consonant confusions indicate that across various signal-to-noise ratios, voiceless fricatives are identified correctly more often than their voiced counterparts.

[35] See also Lulich (2004), who arrives at a similar conclusion concerning Russian [v], also working in a phonetics-based model.

In lack of additional and detailed data on place cues, we, nevertheless, hypothesize that [ʋ̥] in certain positions is (partially) neutralized because the relative perceptibility of its target cues—**voicing** and **noise**—, which are to be perceived simultaneously, is severely curtailed. In accordance with the functionalist stance we briefly introduced in the previous section as well as the aerodynamic premise presented above, we predict that [ʋ̥] in an unfavourably cued context can take two partial neutralization routes. Since noise (turbulence) and voicing cannot simultaneously be maintained in such environments, either (i) noise is preserved and voicing is lost, or (ii) voicing is kept and noise is "sacrificed". In the first scenario—when there is no vocal fold vibration and hence the glottis is open—the aerodynamic premise predicts that a rather noisy devoiced fricative ([ʋ̥̝]) is produced as more air pushes out at a given (and relatively constant) constriction degree, the result of which is considerably more frication. [ʋ̥]'s fricativization under devoicing is in line with the behaviour of other approximants; for instance, [l] and [j] (with wide "constriction") also show frication (thus occur as [ɬ] and [ç]) when they occur in positions where they devoice.[36] In our approach, a necessarily devoiced narrow [ʋ̥] should thus show more friction than devoiced wide approximants. In the second case, when constriction is necessarily less narrow, friction is lost but voicing is maintained, and the sound produced is a labial (wide) approximant, a [ʋ] or [w].

In accordance with Steriade's (1997) hierarchy of environments based on direct reference to the number/quality of acoustic cues to the voicing contrast of obstruents (see (30)), we also claim that the same hierarchy is applicable to the explanation of the neutralization effects of [ʋ̥]. Thus, [ʋ̥]'s cues to voicing (and friction, place, etc.) are best perceptible between two sonorants. The moment [ʋ̥] finds itself before a non-sonorant (an obstruent or silence), its voicing cues are severely reduced in number and quality, and so it is in this environment that the two neutralization routes are predicted (in case the segment is preserved). We assume that languages are free to choose which neutralization strategy they follow: devoicing/strong friction or de-noising/vocalization. The first strategy is selected by languages like Hungarian, Russian; whereas the second one is chosen by Slovak, for instance. Absolute loss of contrast, namely segment deletion, is also a possible option in this model; actually, the

[36] See Padgett (2002, 22) for examples from Norwegian, Iberian Spanish and French. Hungarian [j] also displays a similar conduct in neutralization-prone contexts, such as after an obstruent and before a pause: *lépj* [le:pç] 'step.imp'.

v-deletion strategy was apparent in some of the cases of the experiment we conducted (see section **5.3.2**).

Our predictions as to the (likely) **realization** of Hungarian [ʋ̟] in the various contexts are shown in (31) (we use the same hierarchy of contexts as in (30)):

(31)	context:	V__R	#__R, O__R	V__#	V__O	O__O, O__#, #__O
	realiz.:	[ʋ̟]	[ʋ̟]	[ʏ̟]	[ʏ̟]	[ʏ̟]

Thus, Hungarian [ʋ̟] is licensed only before sonorants (including vowels), in all other contexts, it is likely to devoice and obstruentize. We must note that these predictions on [ʋ̟]'s realizations are founded on the phonetic (aerodynamic/cue-based) premise alone; as we will see, other functional factors (such as the possible coarticulatory effect of the active voicing of a following obstruent) can modify these predictions.

Our cue-based approach can also predict the **phonotactic patterning** of [ʋ̟]. Since its contrast receives the most salient cues before a sonorant, it is these types that we expect to display the most items (thus words with pre-sonorant [ʋ̟]). As we move down on the scale of environments towards those with less and less cues for the recognition of [ʋ̟]'s contrast, we anticipate fewer items to contain this consonant; actually, this model predicts that it is when [ʋ̟]'s phonetic targets cannot be realized (i.e., when it is articulated as [ʏ̟]) that its distribution is restricted. Thus, this analysis predicts what surface variants can appear in the various contexts if a specific set of phonetic targets is to be realized, as well as [ʋ̟]'s gradual phonotactic patterning: both aspects are intimately linked and directly grounded in phonetic factors.

Furthermore, [ʋ̟]'s behaviour with respect to the "**Voicing Requirement**" can also be explained on the phonetic grounds that have been tackled here. In pre-sonorant position, as we argued, [ʋ̟] is realized as a passively voiced (narrow) approximant (in other words, it lacks active/dynamic voicing targets). As such, it is expected to pattern with other sounds that bear similar voicing characteristics, like sonorants. It is for this reason that pre-sonorant [ʋ̟] will not voice a previous obstruent. Obviously, the question fundamentally boils down to the phonetics of active/passive (de)voicing and voicing assimilation. Lack of space prevents us from providing a detailed phonetics-based analysis on why sonorants do not normally actively voice obstruents, and why they appear to nevertheless do so in some languages; suffice it to say that we see the core of the problem to be related to the **coarticulatory** properties of voicing targets

(in this we follow the ideas put forth in Farnetani 1997 and Jansen 2004, among others). Only actively voiced and devoiced sounds are assumed to participate in voicing assimilation as their voicing/devoicing-enhancing gestures can "spill over" into neighbouring segments (mainly those preceding them). Passively voiced sounds, on the other hand, do not possess voice-enhancing gestures, and so they "can have no coarticulatory effect on the voicing control of neighbouring obstruents: [...] there is simply nothing to spill over into flanking sounds" (Jansen *op.cit.*, 108).[37]

When [ʋ̥] is not followed by a sonorant but by a **voiceless obstruent** (as in *savtól* 'acid.abl'), the coarticulation-based voicing assimilation model (together with the cue-driven approach we propose) predicts that [ʋ̥] should appear as a truly voiceless noisy fricative. This is because voiceless obstruents are claimed to be actively devoiced in Hungarian with devoicing gestures that can spill over into [ʋ̥]. Remember that our model predicts that [ʋ̥] should be realized as a devoiced and noisy sound ([v̥]) before obstruents; we assume that the active devoicing gestures of a following voiceless obstruent only enhance the voicelessness of [v̥], and so we expect a sound very close to [f] to be realized, with the consequence of the [f]–[ʋ̥] contrast to be (completely) neutralized. We assume thus that *v* is extremely sensitive to the nature (obstruency, voicing, etc.) of the segment following it (and possibly of that preceding it) because its targeted articulatory gestures are difficult to maintain at the same time, and can only be upheld under optimal circumstances (next to sonorants). Consequently, the model hypothesizes that if *v* is followed by a **voiced consonant**, it will more easily receive voicing from it, and depending on the aperture qualities of that consonant, *v*'s realization will gradually move between stages of (i) a weakly fricated voiced narrow approximant [ʋ̥] (before vowels and sonorant consonants), (ii) a more fricatival voiced

[37] It follows from this line of reasoning that only those languages are expected to display voice assimilation to sonorants where sonorants contrast in active voicing and active devoicing. It is likely that languages that do not contrast voicing in sonorants (thus they are passively/modally voiced) can only show voicing assimilation effects if sonorants target passively (de)voiced/neutralized obstruents, which thus lack targets for voicing (as in the case of word-final neutralized obstruents in Slovak and the Western Transdanubian dialect of Hungarian; see above). As Jansen argues, "if neutralized obstruents indeed lack voicing targets, they should show a greater degree of voicing between a vowel and a following sonorant than actively devoiced obstruents, simply as a result of the **passive continuation of voicing into the constriction phase**. It could well be this increased amount of voicing (relative to utterance-final and [voiceless] contexts) that is interpreted by linguists as voicing assimilation" (*op.cit.*, 119; our emphasis). Further quantitative phonetic research is obviously required here.

narrow approximant/fricative, thus a sound close to [v] (before voiced obstruents or even narrowly produced approximants) to (iii) a very noisy devoiced fricative [ʋ̥/f]. Accordingly, the modified table on the predictions concerning *v*'s probable realizations in Hungarian is (32):[38]

(32)	context:	V__R	#__R, O__R	V__#	V__O	O__O, #__O, O__#
	realiz.:	[ʋ̞]	[ʋ̞] [ʋ̞] or [ʏ̞]	[ʏ̞]	[ʏ̞/f] or [v]	[ʏ̞/f] or [v] [ʏ̞]

In the following section, we will present the results of an acoustic experiment that aimed at checking the various phonetic properties of the realizations of Hungarian *v*, and hence the validity of the predictions of the phonetically-based model we have introduced here.

5.3. Hungarian [v] realizations—an acoustic experiment

5.3.1. Method

In the experiment to be presented here, we focused on the acoustic realization of *v* next to a consonant. We examined *v*C and C*v* clusters word-initially, word-finally and in intervocalic position; the study included V*v*V sequences as well, as a point of reference, since we consider the realization of *v* in this position as the prototypical manifestation of what we described as the narrow labiodental approximant [ʋ̞]. Where it was possible, we only investigated monomorphemic forms (hence to cover the items of Table 1); in some cases, however, we looked at words where a morpheme boundary divided the members of the clusters in question (as in *hívj, hívd* for instance).

Ten native speakers (six female and four male) of Educated Colloquial Hungarian were asked to read out the test sentences (see the Appendix) at a normal speech rate in a sound proof cabin. The age of the speakers were between 22 and 28, with the exception of two subjects, who were aged 58 and 60 respectively. They were not paid for the participation in the experiment. The data were recorded with a Sony ECM-MS907 microphone onto a Sony MDMZ0710 minidisk, digitized at 44100 Hz and resampled at 22050 Hz. The acoustic measurements were computer-analysed using Praat (Boersma–Weenink 2005).

The experiment aimed to measure the following parameters:

[38] We would like to stress again that the scale predicts a **probability** of *v*-realizations (given a normal speech rate/style). The "boundaries" of where the actual realizations may occur are not strictly predicted, and may shift.

(33) (a) spectral moments, primarily **Centre of Gravity (CoG)** (but also skewness and standard deviation)

(b) **voicing** (i.e., glottal pulses)

(c) **formant structure** (transitions from neighbouring segments as well as in the steady state of the segment)

(d) segmental **duration**

Here, we will mostly focus on CoG and voicing parameters.

Spectral moments have been widely used in the literature (Jassem 1979; Forrest et al. 1988; Ladefoged 2003, 156ff; Gordon et al. 2005; Machač–Skarnitzl 2005) to quantify consonantal—in particular fricative—characteristics, especially with the need to distinguish one fricative from the other, across vowel context and speaker.[39] The first spectral moment, CoG (or "centroid"), is a measure for how high the frequencies in a spectrum are on average, it thus represents a spectral mean. CoG is one of the reliable indicators of turbulent noise: the higher the CoG value, the noisier the segment. For instance, for a white noise sampled at 22050 Hz the CoG value is 5512.5 Hz, i.e., half of the Nyquist frequency. It is important to note that there are important language-specific and individual variations in the CoG values of a given fricative; the relative order of the CoG values of different sounds is what is of interest in this paper. The value of (normalized) skewness also indicates how noisy a sound is, it shows how much the shape of the spectrum below the CoG is different from above it, for a white noise skewness is zero. Standard deviation is to be interpreted as a measure for how much the frequencies in a spectrum can deviate from the CoG (for a sine wave, the standard deviation is zero, for a white noise, it is the Nyquist frequency divided by $\sqrt{12}$). In those cases when *v* is realized less fricatival and more approximant-like, the CoG values are less informative, vowel transitions, therefore, the formant structure of the segment are of more interest.[40]

Voicing was simply measured on the basis of periodicity in the waveform as well as the presence/absence of a voice bar in the spectogram.

[39] See, however, Shadle–Mair (1996) and Shadle (2005) on the problems of using spectral moments in reliably quantifying the **differences** between fricatives. They can, however, be reliably used to quantify turbulence from non-turbulence, which is the prime motivation of this paper.

[40] As we have indicated, we primarily focus on the first spectral moment, the CoG, in this paper. Nonetheless, the other spectral moments we examined, skewness and standard deviation, were fully compatible with the CoG values we obtained. For instance, whenever CoG was relatively high, skewness approximated to 0 Hz.

Formant transitions were also checked, in particular for results with low CoG (which thus suggested the presence of a less turbulent segment). Relatively clear formant structure and transitions were indicative of a more vowel-like segment.

Measurement samples were gained the following way. The boundaries of *v* were determined by visual inspection of spectograms and waveforms and listening to the recordings. The selections containing the *v* portions contained "extra" parts of at least 25–25 ms at the edges, as these areas were not used by the analyzer program due to the analysis window shape and size. Thus, the usable area safely included the whole *v* portion. Male voices were band filtered at 0–300 Hz, female voices at 0–400 Hz. Several fast Fourier transform (FFT) spectra were computed for each *v* using a Gaussian window shape with a physical length of 50 ms (which is the sum of the "effective" lengths of 25–25 ms around the cursor in the analysis window). The first measurement was taken at 28 ms, then the 50 ms window was shifted with 5 ms increments until the whole of the segment was covered. In this way several overlapping slices were made of each *v*-segment, depending, of course, on the duration of the given segment sample: for shorter *v*-instances (below 50 ms), a 30 ms Gaussian window was employed (with 15–15 ms around the cursor) and only one or two slices were gained. The multiple windowing method is indispensable in the spectral analysis of sounds (such as fricatives) whose distribution of energy is (quasi-) random, inconstant.[41] Lastly, the CoG, skewness, standard deviation and formant values (F1 through F5) of each slice were computed and finally averaged. Thus, averages were obtained for each *v* occurrence for each speaker, but we also averaged every *v* occurrence for all speakers.

5.3.2. Results

Figure 1 exhibits the averages of the CoG values for all speakers. Axis *y* shows the CoG averages, whereas axis *x* contains all the words we tested in the experiment.

The results of the averages of the CoG values of all test words back up the main hypothesis: *v* occurring in a considerably unfavourable context—for instance word-finally preceded by a consonant—becomes very noisy/fricatival (see the results for *könyv, ölyv, jókedv, nyelv* etc. on

[41] On the problems of the spectral analysis of turbulent sounds, see Ladefoged (2003, 153ff).

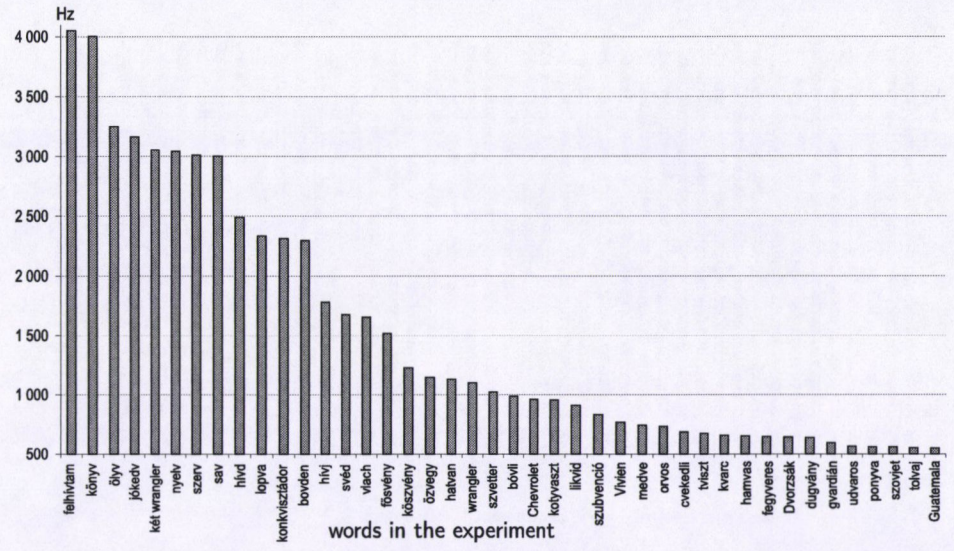


Fig. 1
CoG averages (all speakers)

the left of the figure: the average CoGs of these *v*'s are between 3000–4000 Hz). In a somewhat more favourable context, when *v* is preceded by a vowel but not followed by one, as in *sav* and *hívd*, we still received high CoG values: *v* was realized with a fair amount of fricative noise (2500–3000 Hz). Low CoGs occurred in pre-sonorant position. This is exactly what was indicated concerning *v*'s noisiness in earlier accounts as well (as in Siptár 1996; Siptár–Törkenczy 2000; see (3) on page 183).

It is noteworthy however, and has not yet been noted in the literature to the best of our knowledge, that in cases when *v* is realized as a strongly fricated sound, **it often loses voicing** at the same time. Thus high CoG goes hand in hand with devoicing. This is displayed in Figure 2 (p. 210).

Axis *y* of Figure 2 shows the number of subjects participating in the experiment, axis *x* contains the words tested in a decreasing order of their CoG values. The results back up our hypothesis according to which, *v* can preserve both its friction and voicing targets (as well as place) in pre-sonorant position; in other (unfavourable) positions, it rather tends to lose its voicing and become strongly fricated. Relatively high CoG and (simultaneous) voicing-preservation was observed in pre-obstruent position, where the obstruent is voiced (as in *bovden*). The highest level of CoG and devoicing were found in cases where *v* stands before a voiceless

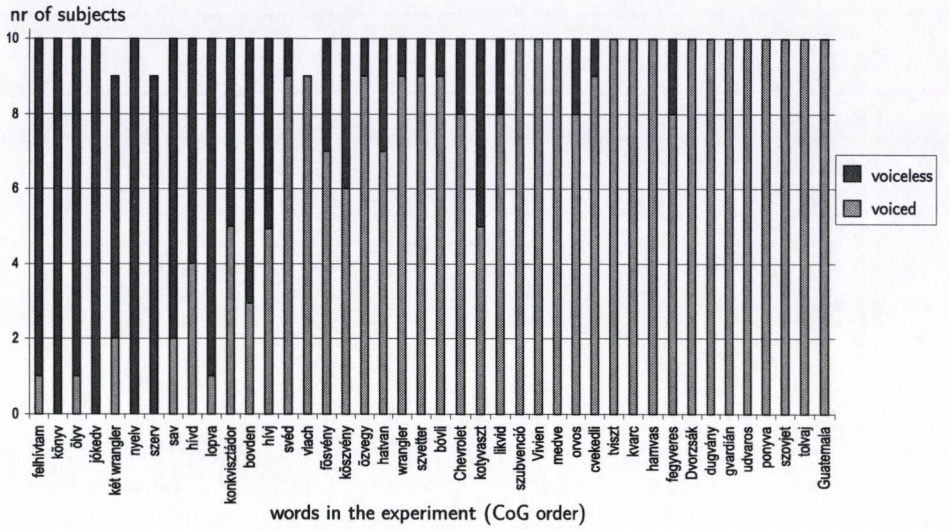
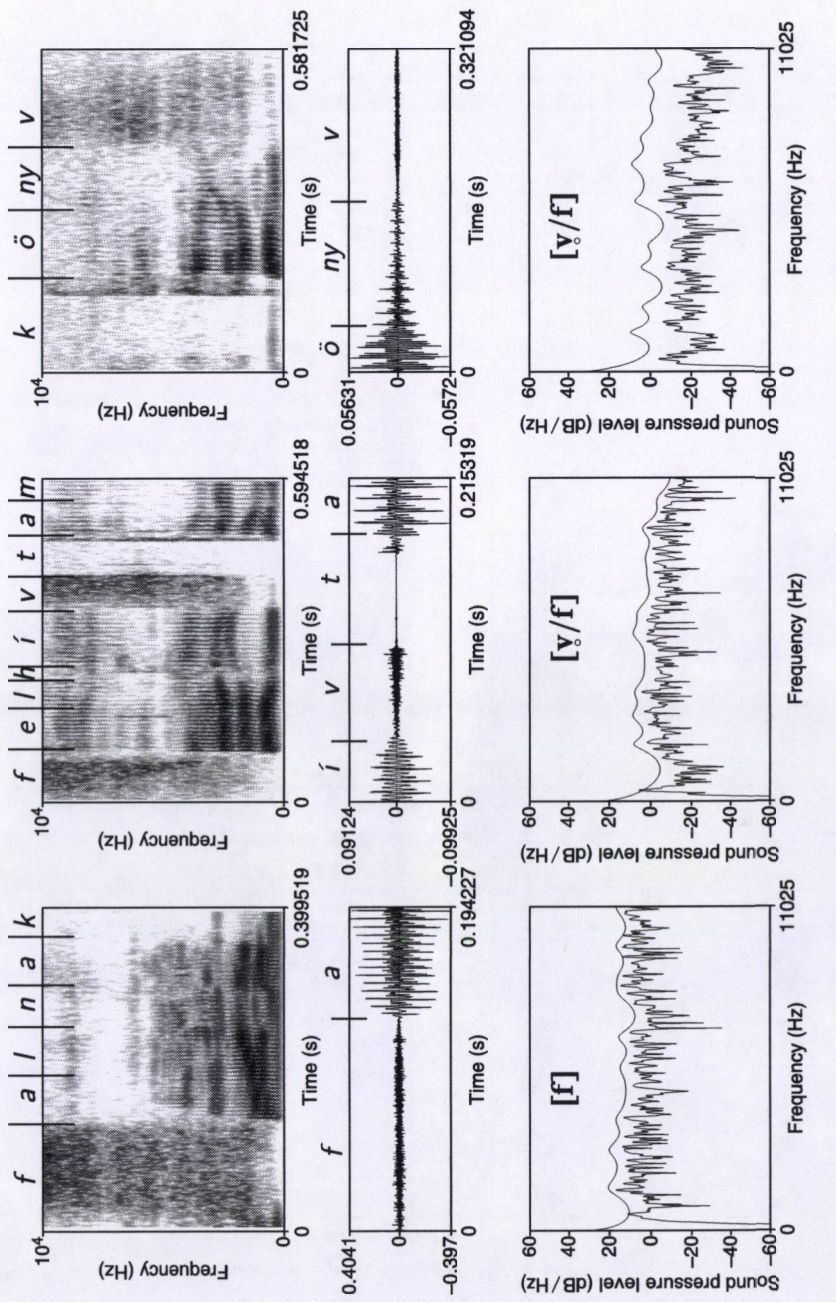


Fig. 2
CoG and voicing

obstruent (as in *felhívtam*), actually, the CoG values of this particular realization is basically the same as that of [f]. Thus, in accordance with the predictions (see (31)), we found the following three-way *v*-realization pattern:

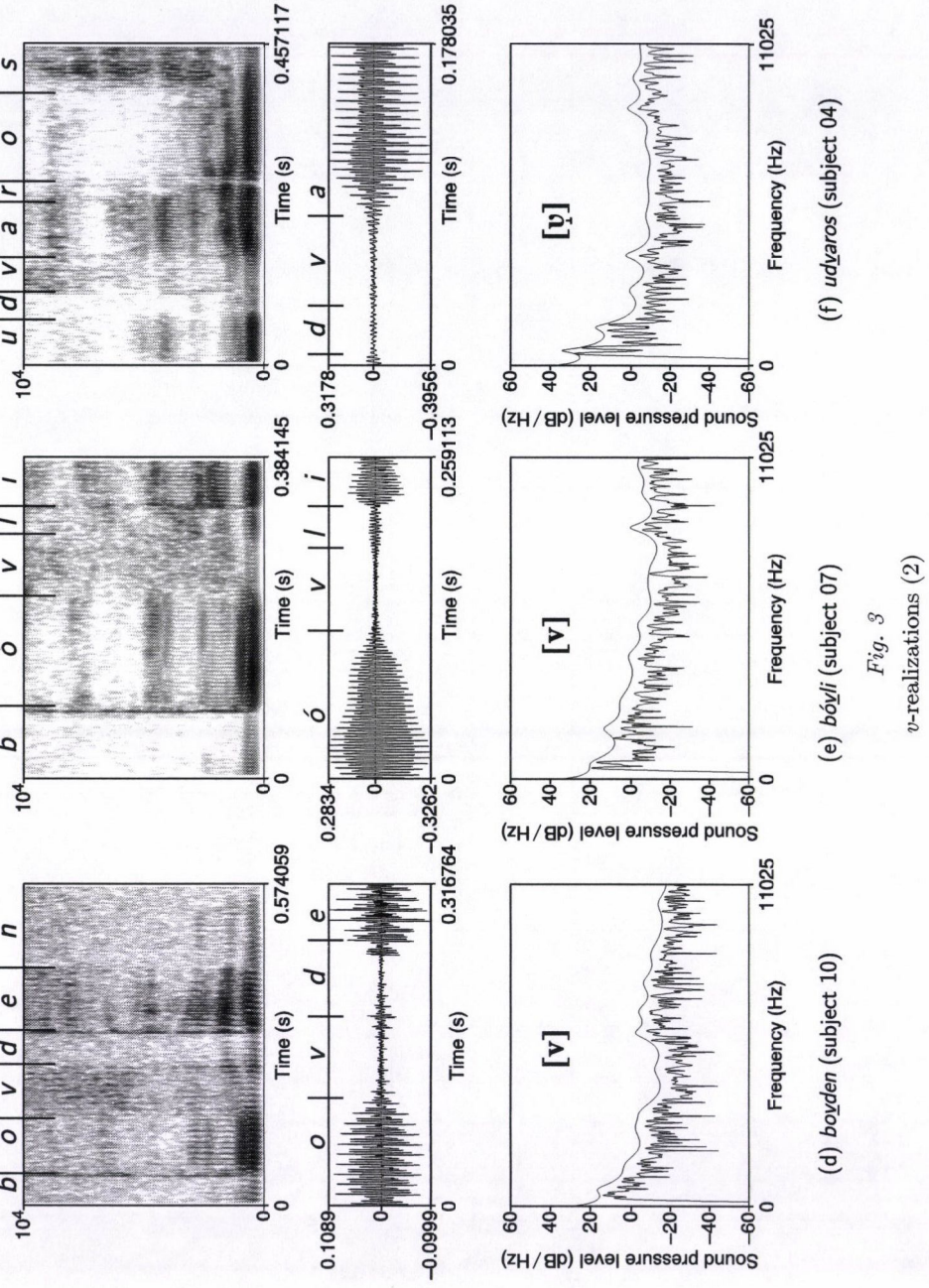
(34) (a) strongly fricated and devoiced [ʋ̥/f] (e.g., *felhívtam, könyv*)
(b) moderately fricated and voiced [v] (e.g., *bovden, bóvli*)
(c) weakly fricated and voiced [ʋ̞] (e.g., *udvaros, Vivien*)

Figure 3 provides spectograms (with corresponding waveforms) and FFT spectra, taken from the middle of the non-transitional portion of *v*, for each realization in (34). Linear predictive code (LPC) smoothed spectra are superimposed in the FFT spectra for easier viewing. We also added an example of a word-initial [f] (*falnak* 'wall.dat') for the sake of better comparison. The similarities between the acoustics of this [f] and the *v*'s of *felhívtam* 'I called' and *könyv* 'book' are striking (see (a–c) on page 211). None of these segments are voiced (see the lack of vertical stripes in the spectograms and the aperiodicity on the waveforms) or contain formant structure (cf. the abrupt start/finish of formant transitions at neighbouring sonorants), and all three show flat/peakless spectra typical of (diffuse)



(a) *falnak* (subject 04) (b) *felhívtam* (subject 01) (c) *könyv* (subject 03)

Fig. 3
f and *v*-realizations (1)



(d) *boyden* (subject 10) (e) *bóyli* (subject 07) (f) *udyaros* (subject 04)

Fig. 3
v-realizations (2)

labial/labiodental voiceless fricatives with energy spread over a large frequency range. We also found some correlation—although the durational aspect of *v* was not the main focus of our study—between length and friction. When *v* was realized as a noisy fricative, it was usually long (around 100 ms), and when it was realized as an approximant, it was usually less than 50 ms long, sometimes even less than 30 ms (this was usually the case in C__V, where the release of the C often masked *v*, and the formant transitions of the V passed through it, as it were).

Examples are given for the *v*'s in *bovden* 'V-shaped belt', *bóvli* 'trash' and *udvaros* 'with a courtyard', respectively in (d–f) (on page 212). Both *bovden* and *bóvli* in these particular instances are voiced but they are also somewhat noisy (not as much, however, as the previous instances of *v* in (a–c)), which can also be seen in their spectra: they are not as diffuse, in fact, lower frequency regions show some intensity. It is these segments that we consider relatively fricated as well as voiced (hence a [v]). The last example (*udvaros*) shows that *v* is voiced (cf. the glottal pulses in the spectogram and the periodicity of the waveform), has a much clearer formant structure continuing into the vowel following it, and its spectrum exhibits characteristic vowel-like peaks at low frequencies. This segment is thus what we described as the narrow labiodental approximant [ʋ̞].

In the "harshest" context (C__#, especially O__#), we always obtained high CoGs and devoicing for *v*, as exactly predicted by the hierarchy of cues in (30), approximant [ʋ̞] was never produced here (one subject deleted *v* in *szerv* 'organ', too). However, this was not the case for the environment V__O: the averages in Figure 1 must be taken with a pinch of salt as they conceal some important generalizations. It is not true that all occurrences of *v* in, e.g., *bovden* were fricated (with an intermediate CoG of around 2000–2500 Hz) and voiced. In fact, and this is evident in Figure 2, seven subjects pronounced *v* here devoiced. Actually, five subjects even devoiced the following *d*, too. Crucially, however, **whenever *v* was devoiced, it was also fricated**, as evidenced by the relatively high CoG averages. The CoG–voicing data of pre-consonantal *v* (as well as of the consonant) for each subject are displayed in Table 5 (overleaf). The data in the table back up our earlier predictions, namely, the exact nature (obstruency/voicing) of the consonant after *v* does influence its realization. It looks as if *v* is most unstable before the voiced obstruent *d* and most stable before *j* ("stable" referring to the fact that *v* can preserve all its articulatory targets). *l* and *r* occupy an intermediate position in affecting *v*. Before *d*, *v* is more likely to occur as

Table 5

Average CoG and voicing of pre-consonantal *v* (1)

bovden				*Chevrolet*				*bóvli*				*szovjet*			
s	CoG	vo-v	vo-C	s	CoG	vo-v	vo-C	s	CoG	vo-v	vo-C	s	CoG	vo-v	vo-C
1	2787	0	1	1	1555	0	0	1	2443	0	1	1	450	1	1
2	896	1	1	2	786	1	1	2	739	1	1	2	627	1	1
3	775	1	1	3	1399	0	1	3	534	1	1	3	497	1	1
4	1992	0	0	4	638	1	1	4	644	1	1	4	692	1	1
5	4619	0	0	5	769	1	1	5	1396	1	1	5	618	1	1
6	3621	0	0	6	2070	0	0	6	1074	1	1	6	708	1	1
7	2904	0	0	7	802	1	1	7	949	1	1	7	486	1	1
8	2372	0	0	8	551	1	1	8	951	1	1	8	594	1	1
9	2194	0	1	9	583	1	1	9	646	1	1	9	483	1	1
10	824	1	1	10	454	1	1	10	508	1	1	10	443	1	1

Legend: "s" = subject; "CoG" = average CoG value (Hz); "vo-v"/ "vo-C" = voicing of *v*/following consonant (0 = voiceless, 1 = voiced)

devoiced/strongly fricated ([v̥]). Only two subjects maintained voicing in the cluster (but then, *v* lost its noisiness by a considerable degree). It is definitely compelling that before *d*, *v* prefers to lose voicing to actually receiving the voicing of the obstruent. We speculate (and consequently, a more thorough research is required) that obstruency is weighted more in the partial neutralization of *v* rather than the coarticulatory spill-over of obstruent voicing: therefore, regardless of *d*'s voicing, this context still counts as partially neutralizing for *v* to a [v̥]. Remember: the markedness of the pre-*d*, and generally pre-obstruent, context is also manifest in phonotactics: such sequences are extremely rare or non-existent. *j*, just like vowels, does not make *v* unstable: here it preserves both its voicing and weak frication, and so it occurs as a [v̞]. With respect to the CoG and voicing parameters, *r* is more like *d*, while *l* is more like *j*. Even though a sonorant, *r* makes *v* lose its voicing (and get fricated) more often than *l*. We assume that this may be due to the trilled articulation of Hungarian *r*, namely, that its articulation involves complete (albeit short and rapid) closures. Nevertheless, we conclude that in this respect, *r* can constitute an unfavourable context for *v*. (Notice that two subjects even devoiced *r* when *v* was devoiced.) Before *l*, the voicing of *v* is preserved; we, however, received a little higher degree of frication than before *j* and vowels. Actually, it was here that we obtained the highest CoG values while voicing was also preserved (see subjects 5 and 6).

Further data also back up our hypothesis that the immediate context of *v* has a great impact on its realization. When the consonant following it is itself in a context marked aerodynamically/perceptually, this impoverished consonant affects *v*, too. Two such cases were measured (*hívd* 'call.2sg.def.imp.', *hívj* 'call.2sg.indef.imp.'), the results are shown in Table 6.

Table 6

Average CoG and voicing of pre-consonantal *v* (2)

	hívd				*hívj*		
s	CoG	vo-v	vo-C	s	CoG	vo-v	vo-C
1	4203	0	0	1	4038	0	0
2	496	1	0	2	2492	0	0
3	1607	1	0	3	2630	0	0
4	1992	1	0	4	762	1	0
5	2950	0	0	5	759	1	0
6	4044	0	0	6	2546	0	1
7	3014	0	0	7	2474	0	0
8	2882	0	0	8	669	1	0
9	840	1	1	9	954	1	0
10	2903	0	0	10	487	1	0

Under such extreme circumstances (i.e., word-finally, after another consonant), *d* and *j* tend to become unstable: both were devoiced in 9/10 cases (*j* actually also became highly fricated when it was devoiced, just like *v*). When these segments were devoiced, *v* was devoiced/fricated in five out of nine cases. When *v* preserved its voicing, its CoG varied, indicating a wide-to-narrow constriction. But whenever we obtained high CoG around 1600–2000 Hz as well as voicing, it involved what we may describe as "careful" pronunciation.

To sum up, we can conclude that the environment C__# constitutes one of the worst contexts for maintaining all the articulatory targets of *v*: it is in this position that *v* is most likely to lose its voicing and become very strongly fricated (in Hungarian at least), **no other alternative is allowed** (disregarding careful speech styles, of course). The environment V__C, on the other hand, is one notch better and thus occupies an intermediate position in the context hierarchy: as we have seen, depending primarily on the characteristics of the following consonant, *v* ([ʋ̞]) can fully be maintained (with low CoG and voicing) **or** lose its voicing and

become fricatival; here thus, due to the intermediacy of the context, the two alternatives that our model predict are indeed possible.

Let us focus now on the **post-obstruent** (and prevocalic) position. This context also showed the same two-fold variation as the pre-consonantal context. *v* either preserved its voicing and was weakly fricated or it was devoiced and rather fricatival. High CoG almost always indicated devoicing, and visa versa. Again, however, the nature of the consonant does seem to play an important role: some consonants are more likely to cause devoicing of *v* than others. For example, in V*tv*V clusters, *v* kept its voicing and weak frication (7/10 subjects) more often than in V*pv*V clusters (3/10). Very similar results were obtained for V*kv*V clusters: in this context, *v* kept its voicing more often. Furthermore, it seems that the fact that the cluster is word-initial is also of importance: word-initially, 10/10 subjects pronounced a voiced *v* with a CoG less than 1000 Hz after *t*, and 9/10 after *k*. Figure 4 shows an example of the two strategies in the pronunciation of *hatvan* 'sixty' (voicing in the spectograms below is also indicated by dotted lines).

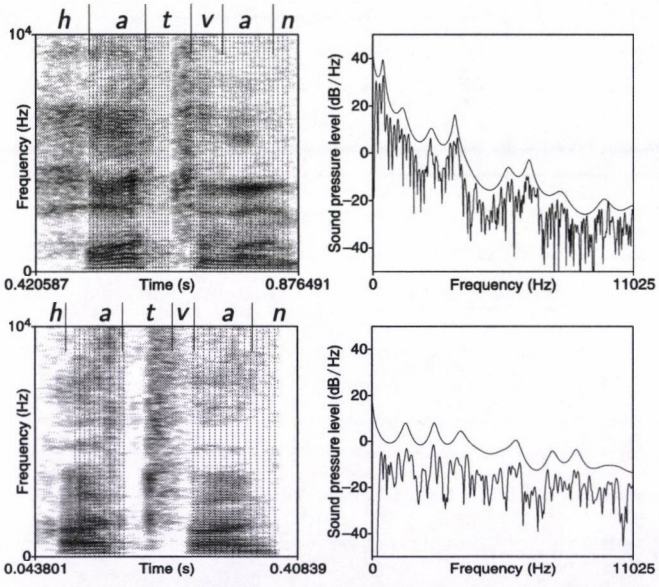


Fig. 4

Realizations of *v* in *hatvan* 'sixty'
(subject 3 vs. subject 10)

The first spectogram and spectra (of subject 3) show a typically approximant-like realization (consider the scarcity of noise at high frequencies, the clear formant structure throughout *v* and the following vowel (*a*), the voicing bars in the spectogram, as well as the formant peaks at low frequencies in the spectrum); the second spectogram and spectra (of subject 10) show that *v* is voiceless, noisy at a wide range of frequencies and lacks well-definable formant structure.

The realizational patterns can thus explain the apparent puzzle of why there is no voicing assimilation in voiceless obstruent–*v* clusters: when *v* is realized as a narrow approximant, its passive voicing cannot spill over to the preceding obstruent, when, however, it devoices and fricates (as in *lopva* 'stealthily'), it being voiceless, there is nothing to voice.

The influence of surrounding consonants on *v* is once again evident in **voiced obstruent–*v* clusters**: in this case, *v* almost never got devoiced, words like *Guatamala, udvaros, dugvány, Dvorzsák, medve, szubvenció* were all pronounced with low CoGs and voiced.

We only noticed devoicing when a **voiceless** obstruent (in some cases *r*, too) stood before *v* and it was pronounced with a relatively strong/long release. The masking effect of the turbulence of release is, we hypothesize, the reason behind the devoicing/frication of *v* in some of the cases. Actually, such masking effects were also observable in [z]*v*, [s]*v* and [ʃ]*v* clusters: here, the fricative noise spilled over into *v* ([z] actually devoiced too in cases like these). We present here two (partial) spectograms showing the two manifestations of *v* in *fösvény* 'miser' as pronounced by subjects 1 and 3 in Figure 5.

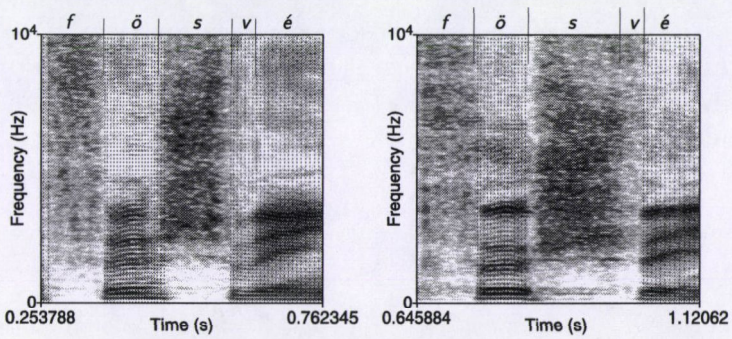


Fig. 5

Realizations of *v* in *fösvény* 'miser'
(subject 1 vs. subject 3)

In the first spectogram, there are unimpeded formant transitions in *v* continuing into the following vowel, *v* is also voiced; in the second, the noise of [ʃ] gradually extends into *v*, the vowel after it begins abruptly; *v* is not voiced here.

We can thus conclude that coarticulation effects play a significant role in the realization of *v* (and hence its phonological behaviour) in C–*v* clusters, too. Further research must find out the exact phonetic motivations. We see place of articulation and its relation to release noise/voicing to exert an important influence on *v*, whose target gestures we argue to be very instable (and fundamentally contradicting) and predisposed to coarticulatory effects. (Lulich's (2004) findings on Russian [v] also appear to support these ideas.)

The realizations of *v* in the remaining set of words, that of **word-initial #*v*C clusters**, also conform to our previous predictions. This context is, again, a highly infelicitous environment to preserve *v*'s targeted noise/voice gestures. This is also shown by the rarity of such clusters in Hungarian (under this analysis an expected result). Their scarcity as well as the fact that they are foreign proper names made the testing rather difficult and therefore conclusive conclusions are hard to draw. It was nonetheless precisely in these tokens that we observed the most variability in the realization of *v* among the subjects. The following four "strategies" were noticeable: (i) some subjects attempted to pronounce these words rather slowly/carefully (as if putting them into "phonological quotes"); in these cases, both voicing and some noise were preserved; (ii) a few subjects pronounced these *v*'s devoiced and with a turbulent noise; (iii) in the case of *Wrangler*, *v* was pronounced by some subjects as if it was an English [w] (even though, of course, no [w] occurs in this word in English), and lastly (iv) we also observed *v*-deletion (hence a complete loss of contrast) in the word *Vlach*. Actually, whenever *v* was preserved as voiced, it was extremely short in all cases (30–35 ms), which made segmentation very difficult.

Two example realizations of *v* in utterance-initial *Vlach* are displayed in Figure 6. In the first spectogram and spectra, we can see a very high degree of turbulence (incidentally, this token showed an average CoG of 4500–5000 Hz) and no voicing whatsoever; the second spectogram and spectra show a very vowel-like realization: notice the distinct and continuous formant structure, the lack of turbulence (average CoG: 559 Hz), and voicing in the spectogram, as well as the intensity peaks at low frequencies in the spectrum. Also take notice of the durational differences between the two *v* realizations.

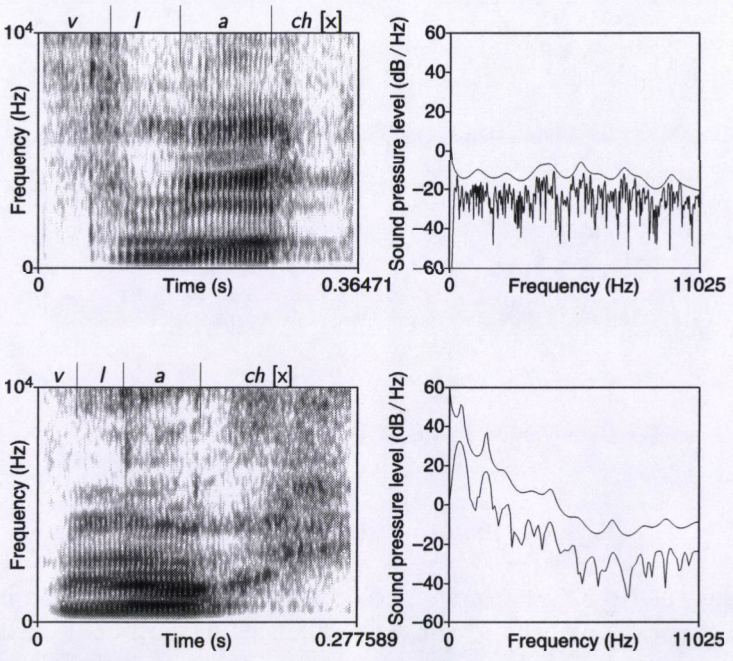


Fig. 6

Realizations of *v* in utterance-initial *Vlach*
'Vlachian/proper name' (subject 3 vs. subject 6)

We also tested *Wrangler* following voiceless *t* in the phrase *két Wrangler* 'two pairs of Wrangler jeans'. The salience of *v* in this position is rather severed; remember: this is actually one of the worst contexts in the cue hierarchy. And considering in particular the obstruent-like manifestation of Hungarian *r*, it comes as no surprise that almost all subjects fricated and devoiced *v* here, sometimes even *r* itself. Here, too, we observed the deletion of *v* with one subject, even when this subject was asked to repeat the test sentence once again. It is thus not unforeseen that assimilation of voicing does not occur: *v* is not actively voiced at all in this position (its is either passively voiced, devoiced or deleted), and so the "Voicing Requirement" is not violated. Figure 7 shows subject 1's pronunciation of *két Wrangler*, in which *v* is realized devoiced/strongly fricated. Clearly, as the lack of glottal pulsing indicates, the cluster [tv̥] is not voiced (actually, even *r* is partially voiceless):

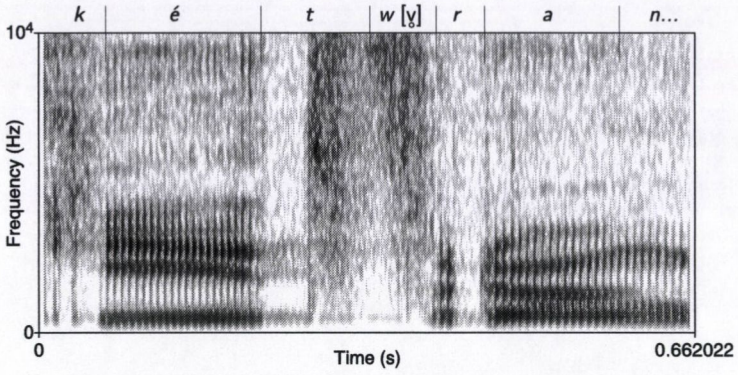


Fig. 7

két Wrangler 'two pairs of Wrangler jeans' (subject 1)

6. Conclusion and remaining issues

The paper has put forth a unified analysis of Hungarian *v*, in which its phonological patterning can be explained in a model based on the phonetic properties of this segment and its linear context. The most important claim has been that the phonetic targets of *v* are inherently contradictory on aerodynamic grounds and can only be maintained in phonetically favourable positions. In such "beneficial" environments, the model predicted the emergence of a passively voiced narrow approximant [ʋ̞]. Segmental contrast of this consonant has been proposed and found to be most salient and hence preservable between sonorants, vowels and (wide) approximants. In other positions, [ʋ̞] is predicted to give up one of its phonetic targets. As a result of this, two realizations are possible: when it devoices, it becomes a strongly fricated, noisy sound; when its voicing target is kept, it loses much of its friction. Hungarian has been shown to be a language which prefers the devoicing strategy in unfavourable positions. As a consequence of the phonetic realizations (owing to specific phonetic factors), its phonological behaviour can also be explained: its role in what we called the "Voicing Requirement" as well as its phonotactic distribution. Neither a prevocalic [ʋ̞], being passively voiced, nor a devoiced [v̥] will induce voicing assimilation as a potential trigger. "Unstable" *v* has been found to be very dependent on the phonetic properties of surrounding sounds. In pre-obstruent position, where it occurs devoiced, active devoicing is easily assimilated by *v*, resulting in a very strongly fricated voiceless sound, which we found very similar

to [f] based on its acoustic parameters. It is thus in this context that the *f–v* contrast is most often lost. If the consonant is voiced after *v*, *v* is realized as a voiced and moderately fricated sound ([v]), this was the case especially before *l*. If, however, a voiced obstruent follows it, *v* is rather devoiced than voiced, therefore, obstruency seems to be a more decisive factor in *v*'s realization than the coarticulation of active voicing. In the surface-oriented model we proposed, thus, all cases of voicing assimilation (including the cases of *lopva/ Wrangler*, having been proven to be problematic for formalist analyses) are explained, without exceptions.

The graduality of *v*'s phonotactics is also accounted for in our cue-based approach. In the least saliently cued contexts, it is predicted that *v* does not occur at all or occurs in only a few items. Most tokens are thus to be found in words in which *v* has several cues of good quality that make its contrast better perceivable. Consequently, in the present analysis, words regarded as "exceptional" in formalist/representational models (like *kedv, hamv*) are as much part of the grammar as the "regular" cases; in fact, their **marginal** status in the phonotactic "space" of the grammar is also predicted. The aerodynamics of *v* and its contexts as well as their cueing potential both predict its distributional regularities, they are intimately linked. *v* is devoiced in unfavourable contexts due to aerodynamic reasons, these contexts are the same as those where cues cannot license its contrast either. It is thus in this way that the current model is directly phonetics-based. Phonetic realizations conduct phonological patterning and the realizations themselves are grounded in phonetic factors.

This analysis is admittedly a first step in a phonetically-rooted phonological analysis of Hungarian *v*, as well as its consonantal phonotactics and voicing/voicing neutralization. Therefore, several issues have had to be left untouched here. Further future research is needed to answer such questions as why certain consonants trigger devoicing of *v* to a greater degree than others (cf. the *tv/kv–pv* asymmetry), how exactly the link between final obstruent devoicing and sonorant voicing can be explained in a phonetics-based model, how the neutralization of *v* can be described (is it partial or complete?), if neutralization is partial, what phonetic parameters help maintain/perceive the contrast, how speakers generally perceive and interpret the devoicing of *v*, how the analysis can be extended to other consonants (especially voiced fricatives and approximants) and to other languages, etc.[42] Even though our aim was not to provide a

[42] *v* does not behave the same way in all languages. For example, in Polish, voiceless obstruents assimilate to *v* and English *v* (just like the rest of the voiced fricatives)

complete account of how a non-formalist phonological framework is built up, we see the current analysis as a useful contribution to the growing body of work on surface-based functionalist phonology, according to which synchronic grammar (or grammar change)[43] is directly influenced by such low-level functional factors as the aerodynamics of articulation, and speech perception.

Appendix

The following test sentences were used in the experiment:

1. Már megint kiment a kocsimban a bovden.
2. Ennyi bóvlit egy rakáson!
3. A szovjet nagykövet új kocsija egy Chevrolet.
4. Hatvan kilométerre ide van egy medve- és farkasrezervátum.
5. Mit kotyvasztasz?
6. (a) A likvid tőke egyszerűen azt jelenti, hogy mennyi pénzed van.
   (b) A mezőgazdasági szubvenciókról többet nem akarok hallani, már unom.
7. A szomszéd özvegyasszony állandóan a dugványaival van elfoglalva.
8. Az a fösvény tolvaj a regényben köszvénytől szenved.
9. Nekem nem tetszett Udvaros Dorottya, egyáltalán nem illett hozzá ez a hamvas ártatlan kislány szerep.
10. A sátorponyva alatt hasaltak a fegyveres őrök.
11. Ne a Lacit hívd!
12. Érted? Engem hívj!
13. Ne aggódjatok! Lesz itt jókedv.
14. Már megint volt az asztalán két új könyv.
15. Az angol is germán nyelv.
16. Nemzetközileg melyik a legértékesebb szerv?
17. Magyarország védett ragadozó madarai közül legritkább a sólyom és az ölyv.
18. Vlach János nyugalmazott vezérőrnagy jelentette fel.
19. Wranglert vett magának, nem Trapper farmert.
20. Két Wranglert vett, nem csak egyet.
21. Városunk nagyjairól festmények készülnek.
22. A spanyol konkvisztádorok havonta két várost döntöttek romba.

---

has also been reported to increase the voicing of a preceding obstruent (Jansen 2004). It seems, however, that the functional factors of contrast and contrast dispersion, in addition to the phonetic factors, play an important role in languages where *v* voices: apparently, *v* only triggers voicing assimilation when it is **actively** voiced (like Polish and English *v* and unlike Hungarian *v*). Preliminary data, however, suggest that if a language has a contrastive actively voiced [v], it also has a contrastive passively voiced approximant counterpart, [ʋ/w]. It is as if active voicing plus friction ([v]) are employed so that the contrast can be maximally salient and distinct from passively voiced and weakly fricated or frictionless [ʋ/w]. Such a two- or three-way contrast is reported for a few languages in Ladefoged–Maddieson (1996) as well as Padgett (2002): crucially, the non-approximant [v] in those languages is actively voiced and fricated. The tentative prediction is thus that we should find the Hungarian way of *v*-patterning in those languages in which *v* does not contrast with other voiced labial/labiodental fricatives/approximants. No doubt, much further research is needed in this area, too.

[43] See the diachronic/evolutionary functionalist approach of Ohala (1981; 1993); Blevins (2004), among others.

23. Tviszt, tviszt, tviszt, itt a kaszinó tviszt!
24. Kvarckristályokat találtak a szomszédos hegyekben, azóta olyan sok a látogató.
25. Dvorzsák-szimfóniákat hallgattunk, attól nyugodtunk meg ennyire.
26. Gvárdián, Guatemala, több nem jut eszembe.
27. Cvekedlit? Úgy érted, káposztás cvekedlit kérek-e? Imádom!
28. Szvettert vegyél, hallod? Talán a falnak beszélek?!
29. Svédcseppeket szedtem a torokgyulladásomra, és két nap alatt elmúlt.
30. Veszprém az egyik legszebb város a Dunántúlon.
31. Vasat találtak és rezet, abból gazdagodtak meg.
32. Véletlen egybeesésnek minősítették az esetet, pedig kizárt, hogy csak erről van szó.
33. Vígan ünnepelték a leszerelést.
34. Anélkül, hogy a fejét megmozdította volna, lopva a tigrisre nézett.
35. Vonattal mentünk Avasig, onnan busszal.
36. Vivien, nehogy megkóstold! Az sav!
37. Vulkánkitöréseket észleltek Indonézia partjainál.
38. Vörös zsebkendőt lengetett, ki tudja, mire gondolt.
39. Tegnap felhívtam az orvost.
40. Pannika már nagyon szeretne óvodás lenni.

References

Balise, R. R. – Randy L. Diehl 1994. Some distributional facts about fricatives and a perceptual explanation. In: Phonetica 51 : 99–110.

Barkaï, Malachi – Julia Horvath 1978. Voicing assimilation and the sonority hierarchy: Evidence from Russian, Hebrew and Hungarian. In: Linguistics 212 : 77–88.

Blaho, Sylvia 2005. Another look at the misbehaving segments of Hungarian voicing assimilation. In: Piñón – Siptár (2005, 35–55).

Blevins, Juliette 2004. Evolutionary phonology: The emergence of sound patterns. Blackwell, Oxford & Malden MA.

Boersma, Paul – David Weenink 2005. Praat: Doing phonetics by computer (Version 4.3.19) [Computer program]. Retrieved July 20, 2005. (http://www.praat.org/)

Chomky, Noam – Morris Halle 1968. The sound pattern of English. Harper & Row, New York.

Cyran, Eugeniusz 1997. Resonence elements in phonology. A study in Munster Irish (PASE Studies and Monographs 3). Wydawnictwo Folium, Lublin.

Cyran, Eugeniusz (ed.) 1998. Structure and interpretation. Studies in phonology. (PASE Studies and Monographs 4). Wydawnictwo Folium, Lublin.

Cyran, Eugeniusz – Morgan Nilsson 1998. The Slavic [w > v] shift: A case for phonological strength. In: Cyran (1998, 89–100).

Farnetani, Edda 1997. Coarticulation and connected speech processes. In: William J. Hardcastle – John Laver (eds): The handbook of phonetic sciences, 371–404. Blackwell, Oxford & Malden MA.

Flemming, Edward 2004. Contrast and perceptual distinctiveness. In: Hayes et al. (2004, 232–76).

Fodor, Katalin 2003. A nyelvjárási hangtani jelenségek [Dialectal phonological phenomena]. In: Kiss (2003, 325–50).

Forrest, Karen – Gary Weismer – Paul Milenkovic – Ronald N. Dougall 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. In: Journal of the Acoustical Society of America 84 : 115–23.

Fuchs, Susanne – Jana Brunner 2005. Conflicting requirements for voiced alveolar fricatives. Paper presented at the Conference on Turbulences, 13–14 October 2005, Berlin.

Gordon, Matthew – Paul Barthmaier – Kathy Sands 2005. A cross-linguistic study of voiceless fricatives. Ms. University of California, Santa Barbara.

Hall, Daniel Currie 2003. Laryngeal feature specifications in West Slavic languages. In: Toronto Working Papers in Linguistics 20 : 93–114.

Hall, Daniel Currie 2004. A formal approach to /v/: Evidence from Czech and Slovak. In: Olga Arnaudova – Wayles Browne – Maria Luisa Rivero – Danijela Stojanovic (eds): Formal approaches to Slavic linguistics 12: The Ottawa Meeting 2003, 187–206. Michigan Slavic Publications, Ann Arbor.

Halle, Morris 1959. The sound pattern of Russian. Mouton, The Hague.

Hayes, Bruce 1984. The phonetics and phonology of Russian voicing assimilation. In: Mark Aronoff – Robert T. Oehrle (eds): Language sound structure, 318–28. MIT Press, Cambridge MA.

Hayes, Bruce – Robert Kirchner – Donca Steriade (eds) 2004. Phonetically based phonology. Cambridge University Press, Cambridge.

Hume, Elizabeth – Keith Johnson 2001. A model of interplay between speech perception and phonology. In: Elizabeth Hume – Keith Johnson (eds): The role of speech perception in phonology, 3–26. Academic Press, New York.

Imre, Samu 1971. A magyar nyelvjárások rendszere [The system of Hungarian dialects]. Akadémiai Kiadó, Budapest.

Jansen, Wouter 2004. Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch. Doctoral dissertation, Rijksuniversiteit Groningen.

Jassem, Wiktor 1979. Classification of fricative spectra using statistical discriminant functions. In: Björn Lindblom – Sven Öhman (eds): Frontiers of speech communication research, 189–206. Academic Press, New York.

Johnson, Keith 2003. Acoustic and auditory phonetics (Second edition). Blackwell, Malden MA & Oxford.

Juhász, Dezső 2003. A magyar nyelvjárások területi egységei [The regional units of Hungarian dialects]. In: Kiss (2003, 262–324).

Kager, René 1999. Optimality theory. Cambridge University Press, Cambridge.

Kenstowicz, Michael – Mahasen Abu-Mansour – Miklós Törkenczy 2003. Two notes on laryngeal licensing. In: Stefan Ploch (ed.): Living on the edge: 27 papers in honour of Jonathan Kaye, 259–82. Mouton de Gruyter, Berlin & New York.

Kiefer, Ferenc (ed.) 1994. Strukturális magyar nyelvtan 2: Fonológia [A structural grammar of Hungarian 2: Phonology]. Akadémiai Kiadó, Budapest.

Kiss, Jenő (ed.) 2003. Magyar dialektológia [Hungarian dialectology]. Osiris Kiadó, Budapest.

Kiss, Zoltán 2005. Graduality and closedness in consonantal phonotactics: A perceptually grounded approach. In: Sylvia Blaho – Luis Vicente – Erik Schoorlemmer (eds): Proceedings of Console XIII, 171–95. Student Organization of Linguistics in Europe (SOLE), Leiden.

Kiss, Zoltán in preparation. The phonotactics of consonant clusters: A perception and articulation-based approach. Doctoral dissertation, Eötvös Loránd University (ELTE), Budapest.

Kornai, András 1994. On Hungarian morphology (Linguistica, Series A: Studia et Dissertationes 14). Linguistics Institute of Hungarian Academy of Sciences, Budapest.

Ladefoged, Peter 2003. Phonetic data analysis: An introduction to fieldwork and instrumental techniques. Blackwell, Malden MA & Oxford.

Ladefoged, Peter – Ian Maddieson 1996. The sounds of the world's languages. Blackwell, Cambridge MA & Oxford.

Lombardi, Linda 1995. Laryngeal neutralization and syllable wellformedness. In: Natural Language and Linguistic Theory 13 : 39–74.

Lombardi, Linda 1999. Positional faithfulness and voicing assimilation in Optimality Theory. In: Natural Language and Linguistic Theory 17 : 267–302.

Lulich, Steven M. 2004. Russian [v]: An acoustic study. In: Folia Linguistica (Acta Societatis Linguisticae Europaeae) 38 : 63–85.

Machač, Pavel – Pavel Skarnitzl 2005. Spectral moments of Czech plosives. Paper presented at the Conference on Turbulences, 13–14 October 2005, Berlin.

Ohala, John J. 1981. The listener as a source of sound change. In: Carrie S. Masek – Roberta A. Hendrik – Mary Frances Miller (eds): Papers from the parasession on language and behaviour (CLS 17), 178–203. Chicago Linguistics Society, Chicago.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In: P. MacNeilage (ed.): The production of speech, 189–216. Springer, New York.

Ohala, John J. 1993. The phonetics of sound change. In: Charles Jones (ed.): Historical linguistics: Problems and perspectives, 237–78. Longman, London.

Olsson, Magnus 1992. Hungarian phonology and morphology (Travaux de l'Institute de Linguistique de Lund 26). Lund University Press, Lund.

Padgett, Jaye 2002. Russian voicing assimilation, final devoicing, and the problem of [v]. Ms. University of California, Santa Cruz.

Petrova, Olga 2003. Sonorants and the labiodental continuant /v/ in Russian voice assimilation: An OT analysis. In: Olga Arnaudova – Wayles Browne – Maria Luisa Rivero – Danijela Stojanovic (eds): Formal approaches to Slavic linguistics 11: The Amherst Meeting 2002, 413–32. Michigan Slavic Publications, Ann Arbor.

Petrova, Olga – Szilárd Szentgyörgyi 2004. /v/ and voice assimilation in Hungarian and Russian. In: Folia Linguistica (Acta Societatis Linguisticae Europaeae) 38 : 87–116.

Piñón, Christopher – Péter Siptár (eds) 2005. Approaches to Hungarian 9 (Papers from the Düsseldorf conference). Akadémiai Kiadó, Budapest.

Rebrus, Péter – Viktor Trón 2002. A fonotaktikai általánosításokról (Kísérlet a magyar mássalhangzó-kapcsolatok nem-reprezentációs leírására) [On phonotactic generalisations (An attempt at the non-representational description of Hungarian consonant clusters)]. In: Márta Maleczki (ed.): A mai magyar nyelv leírásának újabb módszerei V. Szeged, 2001. szeptember 27–29. [Current methods in the description of Hungarian 5. Szeged, 29–29 September 2001], 17–63. Szegedi Tudományegyetem, Általános Nyelvészeti Tanszék, Magyar Nyelvészeti Tanszék, Szeged.

Ritter, Nancy A. 2000. Hungarian voicing assimilation revisited in Head-Driven Phonology. In: Gábor Alberti – István Kenesei (eds): Approaches to Hungarian 7 (Papers from the Pécs conference), 23–49. JATEPress, Szeged.

Shadle, Christine H. 2005. Characterizing fricatives using dynamics of turbulence noise. Paper presented at the Conference on Turbulences, 13–14 October 2005, Berlin.

Shadle, Christine H. – Sheila J. Mair 1996. Quantifying spectral characteristics of fricatives. Paper presented at the Fourth International Conference on Spoken Language Processing (ICSLP96). Manuscript.

Siptár, Péter 1994. A mássalhangzók [Consonants]. In: Kiefer (1994, 183–272).

Siptár, Péter 1996. A Janus-faced Hungarian consonant. In: The Even Yearbook 2: 83–96.

Siptár, Péter – Szilárd Szentgyörgyi 2002. *H* as in Hungarian. In: Acta Linguistica Hungarica 49: 427–56.

Siptár, Péter – Miklós Törkenczy 2000. The phonology of Hungarian. Oxford University Press, Oxford.

Steriade, Donca 1997. Phonetics in phonology: The case of laryngeal neutralization. Manuscript. UCLA.

Steriade, Donca 1999. Alternatives to syllable-based accounts of consonantal phonotactics. In: O. Fujimora – B. Joseph – B. Palek (eds): Proceedings of the 1998 Linguistics and Phonetics Conference, 205–42. The Karolinum Press, Prague.

Stevens, Kenneth N. 1998. Acoustic phonetics. MIT Press, Cambridge MA.

Szentgyörgyi, Szilárd – Péter Siptár 2005. Hungarian *H*-type segments in Optimality Theory. In: Piñón – Siptár (2005, 261–81).

Szigetvári, Péter 1998a. Voice assimilation in Hungarian: The hitches. In: The Even Yearbook 3: 223–36.

Szigetvári, Péter 1998b. Why [h] is not voiced. In: Cyran (1998, 287–301).

Szépe, György 1968. Az alsóbb nyelvi szintek leírása [The description of low linguistic levels]. In: Általános Nyelvészeti Tanulmányok 6: 359–466.

Törkenczy, Miklós 1994. A szótag [The syllable]. In: Kiefer (1994, 273–392).

Törkenczy, Miklós 2000. Lombardi (1999) on voicing assimilation in final clusters in Yiddish-type languages: The embarrassing case of Hungarian. Paper presented at the Massachusetts Institute of Technology, Cambridge, Massachusetts.

Vago, Robert 1980. The sound pattern of Hungarian. Washington D. C., Georgetown University Press.

Wright, Richard 2004. A review of perceptual cues and cue robustness. In: Hayes et al. (2004, 34–57).

Zsigri, Gyula 1994. Magyar mássalhangzószabályok [Rules for Hungarian consonants]. Doctoral dissertation, József Attila University (JATE), Szeged.

INFORMATION FOR CONTRIBUTORS

Please provide typewritten double-spaced manuscripts with wide margins (50 characters per line, 25 lines per page). Type on one side of the paper only and number all the pages consecutively. The following information should be provided on the first page: the title of the article (or, in case of reviews, the title of the book under discussion, the name of the publisher, the place and date of publication as well as the number of pages), the author's full name and address (including email address) and an abbreviated title of the paper (not exceeding 45 characters including spaces). Only original papers will be published and a copy of the *Publishing Agreement* will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement. Authors are requested to send two hard copies of their manuscript + a floppy disk under DOS/Windows. The manuscript should be accompanied by a list of five keywords and an abstract of about 100 words. It should be typed on a separate sheet of paper. Tables, diagrams and illustrations (with the author's name and an Arabic number) should be presented on separate sheets. Captions should be typed on a separate sheet and placed at the end of the manuscript. Footnotes should be typed double-spaced on a separate sheet at the end of the article. All language examples (and only these) are to be italicized (single underlining). Citations in the text should be enclosed in double quotation marks (" " in case of a paper written in English, „ " in German and « » in French).

Reference to a publication should be made by the name of the author, the year of publication and, when necessary, by page numbers in the following ways:

...as described by Schmidt (1967)...
...as referred to by Hoover (1967, 56–78; 1976, 43)...
...mentioned by several authors (Elgar 1978, 77; Williams 1981, 154–6).

An alphabetically arranged list of references should be presented at the end of the article as follows:

Bárczi, Géza 1958a. Magyar hangtörténet [The historical phonology of Hungarian]. Akadémiai Kiadó, Budapest.

Bárczi, Géza 1958b. A szótövek [Word stems]. Akadémiai Kiadó, Budapest.

Brockhaus, Wiebke 1995. Skeletal and suprasegmental structure within Government Phonology. In: Jaques Durand – Francis Katamba (eds): Frontiers in phonology: Atoms, structures, derivations. 180–221. Longman, Harlow.

Cole, Jennifer 1995. The cycle in phonology. In: Goldsmith (1995: 206–44).

Goldsmith, John A. (ed.) 1995. The handbook of phonological theory. Blackwell, Cambridge MA & Oxford.

Kaye, Jonathan – Jean Lowenstamm – Jean-Roger Vergnaud 1990. Constituent structure and government in phonology. In: Phonology 7: 301–30.

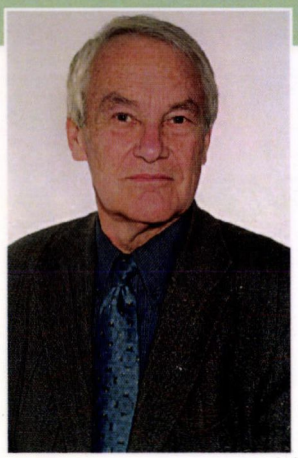
Tomioka, Satoshi 1997. Focusing effects and NP-interpretation in VP-ellipsis. Doctoral dissertation, University of Massachusetts, Amherst.

For marking subsections decimal notation should be applied. Do not use more than four digits if possible. Examples within the text should be marked in italics. Meanings are to be rendered between inverted commas (' '). If glosses are given morpheme by morpheme, the initial letter of the gloss should be placed exactly below that of the example. Grammatical morphemes can be abbreviated in small case letters connected to the stem or the other morphemes by a hyphen. No period should be applied in such abbreviation. For example:

(1) (a) A sólymaid elszálltak
the falcon-gen-pl-2sg away-flew-3pl
'Your falcons have flown away.'

Examples can be referred to in the text as (1a), (1a–d), etc.

One proof will be sent to the author. Please read it carefully and return it by air mail to the editor within one week.



FERENC KIEFER, editor-in-chief, professor emeritus
Member of the Hungarian Academy of Sciences
President of CIPL
Corresponding member of the Austrian Academy of Sciences
Member of Academia Europaea and of
European Academy of Arts, Sciences and Humanities
Main research areas:
morphology, lexical semantics, semantics-pragmatics interface
He has published 11 books (in Hungarian, English, French and German) and more than 200 articles

CODEN: ACTA LINGUIST. HUNG.
SIGILLUM: ALHUE 8

ISSN 1216-8076

phonology ■ morphology ■ syntax ■ semantics and pragmatics ■ sociolinguistics ■ cognitive linguistics

Volume 53 ▪ Number 3 ▪ September 2006

Editor-in-Chief ▪ FERENC KIEFER

FOUNDED IN 1951

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

WWW.AKADEMIAI.COM

AKADÉMIAI KIADÓ

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Acta Linguistica Hungarica publishes papers on general linguistics with particular emphasis on discussions of theoretical issuses concerning Hungarian and other Finno-Ugric languages. Papers presenting empirical material must have strong theoretical implications.
The scope of the journal is not restricted to the traditional research areas of linguistics (phonology, syntax and semantics, both synchronic and diachronic), it also covers other areas such as socio- and psycholinguistics, neurolinguistics, discourse analysis, the philosophy of language, language typology, and formal semantics.
The articles should be written in English, exceptionally, however, papers written in German and French are also accepted. The journal is published in yearly volumes of four issues.

■

Abstracted/indexed in
Bibliographie Linguistique/Linguistic Bibliography, International Bibliographies IBZ and IBR, Linguistics Abstracts, Linguistics and Language Behaviour Abstracts, MLA International Bibliography.

■

Editorial correspondence
ACTA LINGUISTICA HUNGARICA
Research Institute for Linguistics
P.O. Box 701/518, H-1399 Budapest, Hungary
Phone: (+36 1) 351 0413; Fax: (+36 1) 322 9297
E-mail: kiefer@nytud.hu; siptar@nytud.hu

■

Subscription price
for Volume 53 (2006) in 4 issues EUR 252 + VAT (for North America: USD 316) including online access and normal postage; airmail delivery EUR 20 (USD 25).

■

Please send your order to
AKADÉMIAI KIADÓ
Journals Business Centre
P.O. Box 245, H-1519 Budapest, Hungary
Phone: (+36 1) 464 8222; Fax: (+36 1) 464 8221
E-mail: journals@akkrt.hu
www.akademiai.com; www.akademiaikiado.hu

■

ISSN 1216-8076

ALing 53 (2006) 3

Printed in Hungary

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Volume 53, Number 3, September 2006

AKADÉMIAI KIADÓ
MEMBER OF WOLTERS KLUWER GROUP

Cover design: xfer grafikai műhely

CONTENTS

Acta Linguistica Hungarica, Vol. 53 (3), pp. 227–248 (2006)
DOI: 10.1556/ALing.53.2006.3.1

CONDITIONS ON CONDITIONAL MOOD

HUBA BARTOS

Research Institute for Linguistics
Hungarian Academy of Sciences
Benczúr utca 33.
H–1068 Budapest
Hungary
bartos@nytud.hu

Abstract: This paper argues that conditional mood morphology in Hungarian is the spellout of a morphosyntactic feature *excl*(w), the semantic interpretation of which is modal exclusion, i.e., counterfactuality. In certain cases (such as wishes, or CF conditionals) this feature is lexically specified on M[ood], with the direct interpretive aim of counterfactuality, while in others M has this feature unvalued, and inherits its value from the category Mod[ality] in a standard AGREE relation. The strong interrelation between M and Mod also manifests in scope phenomena earlier analysed as scope inversion between Mod and T[ense], but can now be accounted for in a more principled way. Finally, it is shown that, unlike what is found in many other languages, Hungarian cannot use tense marking as the exponent of *excl*(w), because its "tense" is relative, rather than deictic. Therefore, this language makes use of mood morphology to encode CF, in the particular form of conditional mood.

Keywords: mood, modality, tense, counterfactuality, scope

1. Introduction

In modern grammatical theory, the analysis of mood has been a rather understudied subject, and even within this relative poverty, the most neglected of all moods has been conditional mood. While the morphology and phonology of mood marking usually does not bring up very peculiar points (though in the particular case of Hungarian, the morphological

1216–8076/$ 20.00

complementarity of overt mood and tense marking **is** such a point),[1] its syntactic background, and especially its semantic contribution is still a poorly understood area today. This paper aims to shed at least **some** light on the interplay between the morphosyntax and semantics of conditional mood in Hungarian, capitalizing on observations about its apparent scopal behavior (Bartos 2003), and about its semantics (Eszes 2004), as well as Iatridou's (2000) ideas on the notion of counterfactuality.

The leading idea of the paper is that Hungarian conditional mood marking covers a range of various semantic features, and on the level of syntax and semantics, mood, in general, is strongly interlinked with the notion of modality. One of the most obvious roles of conditional mood in this language is to denote counterfactuality, and it plays this role, often assumed by tense marking in other languages, because tense marking in Hungarian is not suitable for this.

2. The data

2.1. Bartos (2003)

Bartos (2003) observed (and gave a syntactic account of) scopal interaction between the three interpretable inflectional morphemes on Hungarian verbs: mood, tense, and modality.[2] His most relevant findings can be summarized like this:

- In a handful of cases we find variable scope order among these inflectional categories, accompanied by rigidly fixed affix order on the verb forms. In particular: a V+Mod+T sequence can be paired with either scope order of T and Mod, though with Mod receiving different interpretations: (1a); a V+Mod+Mood sequence can either be understood as a M > Mod scope order, or as bouletic modality, co-marked, so to say, by modality and mood morphology: (1b); a V+T(+V_{expl})+M sequence can either place tense in the scope of mood, or induce a reading where tense scopes over some modality of volition, whose only indicator is the conditional marker: (1c);

[1] See Antal (1961) for more on this, and Rebrus (2000) and Bartos (2003) for recent theoretical accounts.

[2] In Hungarian, the only overt modality marker is *-hAt*, standing for various flavors of potentiality/possibility; tense can be overt past (*-t(t)*), or unmarked non-past; mood is either conditional (*-nA*), or subjunctive (*-j*), or unmarked indicative.

finally, a form where all three categories are overtly marked (i.e., possibility+past+conditional) has a primary reading with straight scope order (M > T > Mod), but with a secondary reading as well, where M and Mod co-mark bouletic modality again, and standing in either scope order with respect to T (i.e., there is three-way ambiguity here): (1d).

(1) (a) Vár-**hat-t**-ak.[3]
wait-poss-past-3pl
'They could/were allowed to wait.' $T > Mod_{deontic}$
or: 'They may have waited' $Mod_{epistemic} > T$

(b) Vár-**hat-ná**-nak.
wait-poss-cond-3pl
'They could (possibly) wait.' $M > Mod_{deontic}$
or: 'They really should wait.' $M + Mod = Mod_{bouletic}$

(c) Vár-**t**-ak vol-**na**.[4]
wait-past-3pl expl-cond
'They would have waited.' $M > T$
or: 'They fancied waiting.' $T > Mod_{volition}$

(d) Vár-**hat-t**-ak vol-**na**.
wait-poss-past-3pl expl-cond
'They could have waited.' $M > T > Mod_{deontic}$
or: 'They should have waited.' $M + Mod = Mod_{bouletic} <> T$

- These facts pose a problem to standard assumptions about matchings between syntax and morphology, and syntax and semantics (such as the Mirror Principle (Baker 1985), by which we take the syntactic order of the inflectional categories in question to be the mirror image of their morphological order on the verb forms; the syntax of scope (e.g., Aoun–Li 1993); and the hypothesis of a rigid/universal hierarchy of functional projections (most sharply put by Cinque 1999).
- Mood and modality (at least on the level of morphological marking) appear very closely related, sometimes cooperating in marking certain modalities or scope orders.

[3] The following abbreviations are used in the glosses, beside standard ones: cond = conditional mood, expl = expletive V-stem, poss = possibility modality, pv = preverb, def = definite object agreement, ps = possessor.

[4] For an account of the obligatory use of an expletive V-stem (*vol-*) here see Bartos (2003, 40–2).

Eszes (2004) challenges some of these observations from a semantic point of view—he claims that conditional mood marking is not confined to conditional contexts,[5] and in some cases it actually signals counterfactuality (or counterfactual scope relations):

(2) (a) Megnyerheti a versenyt. Eszes (2004, ex. (51))
pv-win-poss-pres.def(3sg) the race-acc
'He may win the race.'

(b) **Megnyerhetné** a versenyt. (*ibid.*, ex. (52))
pv-win-poss-cond.def(3sg) the race-acc
'He might win the race.'

(c) **Megnyerhette volna** a versenyt.
pv-win-poss-past-def(3sg) expl-cond the race-acc
'It was possible for him to win the race, but
this possibility has not been realized.' (*ibid.*, ex. (53))

He argues that (2b) simply pushes the modality expressed in (2a) to a more distant, less likely degree, i.e., the (2a) ~ (2b) difference is much like the *may* ~ *might* difference in English (as analysed by Condoravdi 2001). And in (2c), the role of the mood marking is to signal a forced scope inversion between Mod and T, to T > Mod ('PERF > MOD' in Eszes' terminology), to the effect of expressing counterfactuality.

My goal, therefore, is to propose an account of (i) the semantic contribution of conditional mood, and of (ii) the close interrelatedness of conditional mood and the modalities expressed by the affix *-hAt*, which captures both the M–Mod scope and co-marking effects observed by Bartos (2003), and Eszes' insight on the relation between conditional mood and counterfactuality.

2.2. The various modalities represented by *-hAt*

Before we can proceed with our task, let us survey the various modality types expressed by the affix *-hAt* in Hungarian. (3a–e) are the relevant types from Kiefer (1981; 1985; 2005), while (3f–g) illustrate the distinction drawn between narrow and wide epistemic possibility by Farkas

[5] In fact, when it appears in a conditional context, it must be correlated by another conditional marker in the other clause (antecedent or consequent).

(1992), following Karttunen (1972). Finally, recall that (2c) above is an example of metaphysical modality from Eszes (2004).

(3) (a) Holnap **eshet** az eső. OBJECTIVE EPISTEMIC
tomorrow fall-poss(3sg) the rain
'In view of our knowledge about the world, it is possible that it'll rain tomorrow.'

(b) Péter most az iskolában **lehet**. SUBJECTIVE EPISTEMIC
P. now the school-in be-poss(3sg)
'In view of our knowledge and beliefs about the world, it is quite likely that Peter is in school now.'

(c) Itt **parkolhatsz**. DEONTIC
here park-poss-2sg
'In view of the relevant regulations and norms, you are allowed to park here.'

(d) Innen mindenki **láthatja** a tűzijátékot. CIRCUMSTANTIAL
from.here everyone see-poss-def(3sg) the fireworks-acc
'In view of the circumstances of spacetime, it is possible for everyone to see the fireworks from here.'

(e) Igazán **várhatnál** egy kicsit. BOULETIC
really wait-poss-cond-2sg a little-acc
'Fancy you could wait a little / I'd like you to wait a little.'

(f) Az orvos már **megérkezhetett**. NARROW EPISTEMIC (Farkas 1992, (27))
the doctor already pv-arrive-poss-past(3sg)
'It is compatible with our knowledge that the doctor has arrived already.'

(g) Szép idő van, de **eshetne** az eső. WIDE EPISTEMIC (*ibid.*, (28))
nice weather is but fall-poss-cond(3sg) the rain
'There are some worlds compatible with our knowledge, in which (unlike in our actual world) it is raining.'

Note that the wide epistemic and metaphysical modalities always involve non-indicative mood, and also that they are quite close to each other in many respects. Moreover, Kiefer's subjective and objective epistemic modalities are subcases of Farkas's narrow epistemic type.

2.3. The distribution of conditional mood

In this subsection, some data are presented (partly following the presentation in Kálmán 2001, 166–8) to illustrate that while in many cases conditional mood is (or appears to be) formally selected by some superordinate lexical item, there are various unselected occurrences, as well, so these must be covered by the analysis I wish to propose afterwards.

2.3.1. Conditional mood selected by a higher predicate

The basic cases of conditional mood selection are the following:

- by certain inherently negative predicates (conditional mood optional): (4a);
- some factive verbs under negation, with their factivity canceled (conditional mood optional): (4b);
- non-factives expressing informational uncertainty under negation (conditional mood optional): (4c);
- by various negative items, e.g., negative postpositions (*helyett* 'instead.of', *nélkül* 'without') (conditional mood obligatory with *ahelyett*, optional with *anélkül*): (4d);
- in relative clauses modifying negative quantifiers (conditional mood optional): (4e).

(4) (a) Kétlem/Valószínűtlen/Lehetetlen, hogy Maci Laci elég okos lenne.
doubt-1sg/unlikely/impossible that Yogi Bear enough smart be-cond(3sg)
'I doubt/It's unlikely/impossible that Yogi Bear is smart enough.'

(b) Nem emlékeztek rá/ bizonyították/látszott, hogy
not remember-past-3pl it.sublat/ prove-past-def-3pl/seem-past(3sg) that
Maci Laci elég okos lenne.
Yogi Bear enough smart be-cond(3sg)
'They didn't recall/prove that Y.B. was smart enough./It didn't look like Yogi Bear was smart enough.'

(c) Nem valószínű/feltételezzük/tartunk attól, hogy Maci Laci túl okos lenne.
not likely/suppose-1pl/keep-1pl it-from that Yogi Bear too smart be-cond(3sg)
'It is not likely/We don't suppose/We are not afraid that Yogi Bear is too smart.'

(d) ... ahelyett/anélkül, hogy aludt volna
that-instead.of/that-without that sleep-past(3sg) expl-cond
'instead of / without (him/her) sleeping'

(e) [Nem kérek tőled] semmi (olyat), ami nehéz lenne.
not ask(pres)1sg from.you nothing such-acc what difficult be-cond(3sg)
'[I'm asking you for] nothing that would be difficult.'

While it is certainly possible that these constructions are reasonably analysed via (optional or obligatory) selection for the feature COND by the relevant superordinate items, in the light of the subsequent examples and discussion we will want to reconsider the issue in section **3**.

2.3.2. Unselected conditional mood

We most frequently find unselected occurrences of conditional marking in

- irrealis conditionals (both in the antecedent and in the consequent): (5a, b);
- in wishes: (5c, d);
- indicating counterfactuality with certain modalities (metaphysical, wide epistemic; cf. section **2.2** above): (5e);
- marking bouletic or volitional modality: seen in ((1b, d), (2e)) and (1d), above, respectively.

(5) (a) Ha **esne** az eső, **hazaindulnánk**.
if fall-cond(3sg) the rain home-start-cond-1pl
'If it rained, we'd be heading for home.'

(b) Ha **esett volna** az eső, **hazaindultunk volna**.
if fall-past expl-cond the rain home-start-past-1pl expl-cond
'If it had rained, we would have been heading for home.'

(c) Bárcsak **havazna**!
if.only snow-cond(3sg)
'If only it snowed!'

(d) Hogy **rohadna** el!
that rot-cond(3sg) away
'Damn!' [lit.: '(Wish) it would rot!]

(e) Szerintem Maci Laci **megnyerte volna** ezt a versenyt.
according.to-1sg Yogi Bear pv-win-past(3sg) expl-cond this-acc the race-acc
'I think Y. B. would have won this race [had it not been for some circumstance that rendered it otherwise].'

Furthermore, it is interesting to note that in earlier stages of Hungarian, the conditional mood was used in subjunctive-like function in subordinate clauses of purpose and time:

(6) (a) %"Szintén immár lóra ugrani akara,
also now horse-onto jump-inf want-past.3sg

Hogy Arszlán táborához **elszaladhatna**"[6]
that Arslan camp-3sg.ps-to away-run-poss-cond(3sg)
'Also now he wanted to jump on a horse // so that he could run to Arslan's camp.'

(b) %[Egy nyavalyás kicsiny bárány budosóba esék.] Melyet mikoron
a wretched little lamb hiding-into fall-past.3sg which-acc when

az farkas **megtalált volna**, nem akará mindjárást megenni [...][7]
the wolf pv-find-perf expl-cond not want-past.3sg at.once pv-eat-inf
'[A wretched little lamb went into hiding.] When the wolf found her, he didn't want to eat her immediately.'

While not directly relevant to our account, these historical data are nevertheless suggestive of our analysis being on the right track, as will be seen presently.

3. The semantic contribution of conditional morphology

Let us now take a look at what we find in the literature on the ways of expressing counterfactuality (CF). The most directly useful piece is Iatridou (2000), where the emergence of CF (as in wishes and non-realizing conditionals) is attributed to a morphosyntactic feature *excl*(F), where F ranges over times t and worlds w: temporal or modal exclusion ("The utterance time/world is excluded from the topic time/world"). The former, *excl*(t), is interpreted as tense, i.e., time relations, the latter yields modal relations, and is a key ingredient of CF—in fact, the meaning of a CF conditional is none else than the meaning of the corresponding non-CF conditional plus the statement "the topic worlds exclude the actual world" (*op.cit.*, 245–7). And a general picture that emerges from a cross-linguistic survey is that most often, what encodes CF is the same as what encodes *excl*(t), e.g., tense morphology—as is the case in English:

(7) (a) If Vic invited me, I **would** go.
(b) If Vic **had** invited me, I **would** have gone.
(c) I wish Vic (had) invited me.

[6] *Szigeti veszedelem* II/14. [*The peril of Sziget*, ch. 2, verse 14], by Miklós Zrínyi (17th c.).

[7] *Az farkasról és bárányról* [*Of the wolf and lamb*], by Gábor Pesti (16th c.)—cited from É. Kiss (2004), and the glosses reflect her analysis in certain respects.

(7a) contains "fake past": it is a statement holding at the time of utterance, so *excl*(F) is not *excl*(t), but *excl*(w), though, of course, in many other instances past (or, to put it more aptly: *excl*) morphology gets interpreted as *excl*(t). The antecedent of (7b) contains two "layers" of *excl*(F), one for t, and another for w, hence the use of the past perfect. CF wishes, as (7c), show a similar effect.

In languages which appear to encode *excl*(F) with subjunctive mood, it is always the **past** subjunctive that is used, i.e., the encoder is, again, some special past tense. 'Conditional mood' as such does not exist, Iatridou claims: what is labeled as 'conditional' (e.g., in French) is actually indicative mood augmented with special tense/aspect. In Hungarian, however, there is little reason to think so,[8] in my opinion, so at this point I diverge from Iatridou's argumentation.

With this reservation, let us check the Hungarian data. Obviously, in all the relevant Hungarian examples (cf. (5)), we find conditional mood marking, in wishes, as well as in both clauses of conditionals, while the tense marking simply serves to locate the event/proposition in time, past serving the *excl*(t) purpose. From a functional perspective it is therefore straightforward to conclude that Hungarian conditional marking plays the same role of indicating CF as past morphology in English, and many other languages. But what is even more significant is that if we look at the 'selected' instances of conditional mood in (3), we find that they, too, appear in counterfactual contexts, or at least non-factual contexts where the possibility of the realization of the event is very slight (cf. the cases of "future less vivid" in Iatridou (2000), which share the counterfactual morphology in various languages). The only quirk is the optionality of conditional morphology here, but even this falls in place at least as well as in a selection-based analysis: if it is meant to signal CF, then it is somewhat redundant, insofar as the matrix domain makes it clear that the embedded proposition is CF, i.e., *excl*(w)-marking on the embedded predicate is reinforcement, rather than primary clue. In a selectional analysis, on the other hand, optionality is but a diacritic on the selectional feature. There is one potentially worrisome circumstance though: in some cases the optional conditional mood alternates with the indicative, while in others, the subjunctive:

[8] See the Appendix, though, for some speculation on this point.

(4) (a′) Kétlem, hogy Maci Laci elég okos.
doubt-1sg that Yogi Bear enough smart
'I doubt that Yogi Bear is smart enough.' INDICATIVE (unmarked)

(a″) Valószínűtlen/Lehetetlen, hogy Maci Laci elég okos legyen.
unlikely/impossible that Yogi Bear enough smart be-subj.3sg
'It's unlikely/impossible that Yogi Bear is smart enough. SUBJUNCTIVE

The (4b, c)-types go with the indicative, alternatively. (4d, e) are only slightly different: in past tense, conditional morphology is obligatory, while it alternates freely with the indicative in the present.

A thorough treatment of mood choice in these constructions would go beyond the scope of the present paper, but following the lead of Farkas (1992), Rákosi (2001), and Tóth (2006), we must suppose that the modal semantics of the matrix determines the choice: the less compatible the topic world of the embedded proposition with the real world of the utterance, the less realis the mood —indicative being more realis than subjunctive, which, in turn, is more realis than the conditional. In the Kratzerian parlance of Tóth, the more remote the worlds in which the embedded proposition hold from the actual world, the less realistic the modal base, and the less realis the mood.

Coming back now to our main concern: recall Eszes' (2004) proposal: in the metaphysical modality cases (e.g., (2c)), the conditional morphology serves to indicate "CF scope order" between Mod and T (T > Mod). This, again, ties in with the above-made observation about the role of conditional mood marking: in Hungarian, the marker of *excl*(w), i.e., of counterfactuality, is obviously none else than conditional mood morphology. Furthermore, this observation paired with the one about "fake" conditional mood expressing some sort of modality (volition-disposition as in (1c)), or combining with Mod-marking to express some other sort of modality (bouletic, as in (1b, d)) leads us to investigating the formal morphosyntactic relation between Mood (M) and Modality (Mod) in Hungarian, as laid out in the next section.

4. Conditional mood and modal possibility —An AGREE relation

Summarizing the relevant content of the data sections above, we have identified the following cooccurrence facts between COND and POSS:

(8) (a) **Modalities cooccurring with conditional mood:**
- wide epistemic
- metaphysical
- volitional-dispositional (with semantically apparently empty M_{cond})
- bouletic

(b) **Modality cooccurring with indicative mood:**
- narrow epistemic

(c) **Modalities cooccurring with any mood:**
- deontic
- circumstantial

I assume with Bartos (2003), in keeping with the Mirror Principle and the syntactic representation of scope relations, the following projectional hierarchy: [M [T [Mod [...]]]].[9] Under standard assumptions, this means that M and Mod cannot select one another, since they are not in the local relationship necessary for that kind of relation, because of the intervening T. We thus conclude that these cooccurrences are best accounted for in terms of an AGREE relation, in a Chomsky (2001)-type syntactic framework. In particular, as regards the system of the relevant formal features, let us adopt Kratzer's (1981; 1991) theory of modality, with the following feature breakdown:

(9)

modal force	→	[± poss]	+ poss → ep, deont, circ, boul, ...[10]
			− poss → vol
modal base	→	[± ep]	+ ep → ep
			− ep → deont, circ, boul, metaph, ...
ordering source	→	[*x* src]	*x* = {deont, boul, stereotypical, ...}
modal exclusion (*excl*(*w*))	→	[± **excl**]	+ excl → vol, metaph, w-ep, ...
			− excl → n-ep, circ, deont, ...

That is, we have the three main ingredients of Kratzer's theory: modal force, modal base, and ordering source, represented by one formal feature each: possibility, epistemicity, and source, respectively. The various

[9] For detailed arguments, and for an account of why/how the expletive V-stem *vol-* appears when T = past and M = cond, see Bartos (1999; 2003).

[10] The abbreviations: ep = epistemic, circ = circumstantial, deont = deontic, boul = bouletic, vol = volitional, metaph = metaphysical, n-ep = narrow epistemic, w-ep = wide epistemic.

available values (binary for the former two, and multiple for the last one) characterize the different modalities as shown in (9). This system is completed by the addition of the modal exclusion feature, as an instantiation of the more general *excl*(F). In sum: these features define the various modalities of the category Mod, and relate it to the category M via the AGREE relation. The key element in capturing the cooccurrence effects is [±excl]. In this respect, narrow epistemic, circumstantial, and deontic modalities display no modal exclusion/distancing: the proposition in their case is potentially valid in the actual world. The other types, on the other hand, do involve such distancing, hence the specification [+excl], so they can be characterized as CF, in Iatridou's sense, whereby conditional morphology will possibly (co)occur with them.

The most interesting point in (9) is the split between narrow and wide epistemic modality with respect to the [excl] feature, and this split follows from a distinction within the epistemic modal base: if the modal base is totally realistic (in Kratzer's (1981) sense), i.e., the real world is inside the set of worlds in which the proposition is claimed to be realized, as is the case for narrow epistemics, the feature is specified for non-exclusion, while if the modal base is only partially realistic, i.e., the real world is not necessarily in the anchor set, as in the case of wide epistemics (Farkas 1992), [+excl] is the appropriate specification.

The concrete mechanism is the following: M may either have an independently specified [excl] feature (e.g., when so selected by some higher head, such as C in conditional clauses, or by some covert item ('I wish ...'), as in wishes), or it may inherit the value by agreeing with Mod. [+excl] is then spelled out as conditional mood marking, and [−excl] as indicative mood marking. In the following subsections we will see how this works in the various cases.

4.1. Conditional marking, no *-hAt*

4.1.1. Wishes and conditionals

On one hand, it is tempting to assume that wishes contain an overt or covert embedding speech-act predicate ('I wish ...'), and/or a particular Force^0 (*bárcsak,* as in (5c) above), which exert selection for a [+excl] feature on M. That is, in the case of wishes the emergence of conditional morphology could then be put down to pure formal feature selection. However, it is quite clear that (i) these sentences are paradigmatic cases

of self-contained counterfactuality, i.e., they can easily be analysed with their M lexically specified as [+excl], directly towards a CF interpretation; and that (ii) a hypothetical matrix predicate with the 'I wish' meaning in Hungarian, such as *azt kívánom ...* 'I wish...' or *(azt) szeretném* 'I would like [it to be the case] that ...', when overtly expressed, selects for the subjunctive either obligatorily (as the former), or optionally, besides the conditional (as the latter), unless *bárcsak* is inserted, which guarantees the conditional morphology:

(10) (a) Azt kívánom, hogy álljon el az eső.
it-acc wish-(pres)1sg that stop-subj-3sg pv the rain
'I wish it stopped raining./I want for it to stop raining.'

(a') %Azt kívánom, hogy bárcsak elállna az eső.
it-acc wish-(pres)1sg that if.only pv-stop-cond(3sg) the rain
'I wish it stopped raining./ I want for it to stop raining.'

(b) Szeretném, hogy elálljon az eső.
like-cond-def-1sg that pv-stop-subj-3sg the rain
'I want it to stop raining.'

(c) Szeretném, ha elállna az eső.
like-cond-def-1sg if pv-stop-cond(3sg) the rain
'I'd like it to be the case that it stops raining.'

That is, it is neither necessary nor adequate to appeal to formal selection — we must take the conditional in non-overtly-embedded wishes as an independently motivated denoter of CF. (Further arguments against attributing the CF nature of wishes to a hypothesized, covert matrix predicate is found in Iatridou (2000, 243).)

In conditionals, CF may affect the antecedent (it provides a condition that holds in worlds incompatible with what is the case in, or what is known of, the actual world) — thus conditional mood as the *excl*(w) marker in such clauses occurs in its own right, again, with the semantic consequence of CF interpretation. As regards the consequents of conditionals, they most probably match the antecedent in Excl-marking,[11]

[11] Iatridou (2000, 268) hints that this may be via some sort of agreement relation between the two clauses, but notes, qutoing Cho (1997), that there are languages where only the antecedents are marked for CF, but not the consequents,

but often have a futurity component,[12] too, and the same CF morphology often appears on the 'wish' matrix verb of CF wishes if it is overt—in Hungarian the interclausal matching holds, and at least one 'wish' V (*szeretné, ha ...* '(s)he would like it if ...') does bear CF-morphology, but there is no sign of futurity whatsoever.

4.1.2. Other allegedly selected cases of conditional mood

In the examples of (4) above, what embeds the clauses marked for conditional mood are contexts which negatively evaluate the (probability of the) truth of the embedded proposition: inherent negatives, negated factives, etc. As pointed out by Farkas (1992, 220), in such contexts, mood choice of the embedded clause depends on the modal distance of the topic world from the actual world (i.e., how realistic the modal base is). Augmenting her basic distinction between indicative and subjunctive by adding the conditional as an 'in between' case, we arrive at the following realizations:

- very close or identical worlds / totally realistic modal base → indicative
- more remote worlds / partially realistic modal base → conditional
- remotest worlds / incompatibility with what is known of the real world → subjunctive

As seen in the examples, too, negated categorical epistemic predicates favor the subjunctive, while non-categorically evaluative ones settle for the conditional—in this respect, both conditional and subjunctive indicate counterfactuality.

The conclusion is that in these cases we need not (and in fact, should not) posit mood selection by a matrix head; instead: semantic compatibility is the decisive factor in determining mood marking in the embedded domain.

4.1.3. Volition/disposition

This is a true instance of specifying the value of M via agreement with Mod. The details are these: Mod bears the features [− poss, + excl]

i.e., marking on the consequent is not universally necessary. Hungarian displays agreement/matching between the clauses, though.

[12] In fact, Iatridou (2000) claims that the so-called conditional mood in many languages (e.g., French) is really a compound of an Excl-marker and a future-marker.

the spellout of which is ∅ (zero morphology), and the interpretation it is assigned is volition or disposition, cf. (1c) in section **2.1.** M, on the other hand, bears an unvalued feature [*u* excl:], which must be valued through an AGREE relation with the relevant feature on Mod: [excl: +]. Once M has this feature valued as '+', it will be spelled out as the suffix *-nA* (and the CF in interpretation concomitantly obtains). Note, though, that some archaic and dialectal forms (11a), and an idiomatic nominalizational pattern (11b), still retain overt *-hAt* for this kind of modality, while the productive pattern displays the zero marking:

(11) (a) %Alhatnék/Alhatnám.
sleep-poss-cond-1sg
'I'd like to sleep.'

(b) Alhatnékja van.
sleep-poss-cond-1sg-3sg.ps be-pres.3sg
'He'd like to sleep.' [lit.: 'He has sleep-wish.']

4.2. Conditional mood in cooccurrence with *-hAt*

4.2.1. Deontic and circumstantial modalities

In these cases Mod is [−excl] (see (9)), so unless M is marked otherwise (i.e., unless it is CF by some other means), it gets its relevant feature valued as in the AGREE relation with Mod [−excl], yielding indicative mood morphology on the spellout side.

4.2.2. Narrow epistemic modality

Let us now turn to the more complicated case of epistemic modality. Here we are immediately faced with a distinction between narrow and wide epistemic modality, in the parlance of Farkas (1992). There appear to be two hallmarks of narrow epistemic modality: (i) it only occurs with indicative/unmarked mood, and (ii) it takes scope over tense (see (1a), (3f) above).

Concerning (i), we might choose to follow Farkas' solution of semantic compatibility (*op.cit.*, 220), but this alone would not yield (ii). Let us therefore see if there is any explanation covering both (i) and (ii). As regards the scope order, we seem to have two options (in the wake of the discussion in Bartos 2003). Either (a) we posit a category $\text{Mod}_{\text{n-ep}}$ in a position c-commanding T in syntax, as distinct from other instantiations

of the category Mod below T, but then we lose any straightforward account of the single identical affix order corresponding to both, or (b) we argue that the scope effect is only apparent. Bartos (2003) chose the former option, and was forced into a not particularly explanatory analysis of the scope switch. Here we, on the other hand, opt for devising a treatment in the latter vein. Let us assume that (against all appearances) the morphologically unmarked M in this case is **not** indicative but **evidential** (cf. a similar suggestion for German in Drubig 2001, pointing back to ideas presented by Westmoreland 1995). M can thus be specified for an [evid] feature, the value of which matches the value of Mod's [ep]-feature, i.e., M will be specified as [− excl, + evid],[13] which is realized as ∅, and interpreted as 'in view of our experience/perception of facts of the world it is possible that *p*'. In essence, therefore, what scopes over T is not Mod itself, but M_{evid}, which is morphologically homonymic with M_{ind}, but constitutes a separate mood. While it would be rather difficult to find any direct, surface evidence for such a mood distinction in Hungarian, the scope "reversal" effect attested in (1a) can be seen as covert, secondary indication of the validity of this distinction, nevertheless.

4.2.3. Wide epistemic modality

Next, we must tackle the wide epistemic modality type. Unlike narrow epistemic modality, it (i) cooccurs with M_{cond}, and (ii) scopes below T. Since such Mod has the values [+ poss, + ep, + excl], M will get valued as [+ excl] via AGREE, yielding the spellout *-nA*. Mod has no access to any position scoping over T, and M is not evidential in this case: we are not considering our experience or perception of facts of any world, but flash up hypothetical possibilities pertaining to "worlds that are not compatible with what is known about the actual world" (Farkas 1992, 220).[14] The absence of "inverse scope" effects between Mod and T thus follows, too.

4.2.4. Metaphysical possibility

The final case that needs to be considered is metaphysical possibility. Mod in this case is specified as [+ excl, + poss], but [− ep] (the modal

[13] Whether M is [− excl] by virtue of its own [+ evid] feature, or it is valued that way by Mod is immaterial at the moment.

[14] Note that Farkas explicitly labels wide epistemic possibility as CF, which ties in with the crucial insight of the present analysis: CF is the interpretation of [+excl] on M.

base is non-epistemic), therefore it will value M with [+excl], resulting in conditional morphology and CF interpretation, just as expected. Disregarding the epistemicity difference, this modality behaves analoguously to wide epistemic possibility, as far as its relation to mood is concerned.

To sum up the results of this section: the various occurrences of conditional morphology have been shown to be the reflex of the feature [+excl] on the category M in the structure, which is either due to some independent factor (selected/forced CF interpretation by the matrix context), or conferred on M in an AGREE relation with Mod, which may bear this feature by lexical specification.[15]

5. Why not T?

In this final section, we seek the answer to the question of why it is M, rather than T, that assumes the role of encoding CF in Hungarian, unlike in many languages, including English, where CF is an alternative interpretation of what is usually termed 'past tense' (Iatridou 2000). Given that it is this "fake" T that is the most (proto-?)typical encoder of *excl(x)*, why does Hungarian opt for using conditional mood to this end?

There seem to be two available answers to this question.

- Answer 1: Because tense and mood are "faces of the same coin", in the sense of Antal (1961), so they are functionally equivalent, and because M is higher in the clause structure than T, CF is marked on M. However, the argument about the identity of mood and tense is narrowly morphological, and does not hold at the level of syntax. While it is certainly true that (i) the subjunctive is tenseless, that (ii) the unmarked "combination" of nonpast tense and indicative mood is plausibly seen as a single zero, instead of two zero items in a sequence, and that (iii) simultaneous past tense and conditional mood can only be expressed by resorting to a complex V-form (cf. (1c), and the discussion in Bartos 2003), the possibility of syntactic **and** semantic cooccurrence of past and conditional takes away most of the motivation for this answer – it is precisely these two that are

[15] It is not hopeless to actually derive the value of [excl] on Mod from the other modal specifications, but whether this is the correct view remains to be explored by future research. Here I take this feature to be lexically valued on Mod, for the time being.

the candidates for marking CF, so if they can cooccur, it is not clear why one would be more suitable for this task than the other.

- Answer 2: Because morphological tense is not deictic tense in Hungarian, but relative tense (or aspect), which is less appropriate for expressing the notion of exclusion. The arguments for the non-deictic nature of tense in Hungarian are the following:
  - T is not deictic, but relative: there is no direct matching with real-world temporality (12a, b), and no sequence of tenses effect in this language (12c):

(12) (a) **Holnap** ilyenkorra már átlép**tük** az Egyenlítőt.
tomorrow by.this.time already over-step-past-1pl the Equator-acc
'By this time tomorrow, we will already have crossed the Equator.'

(b) Eredetileg úgy volt, hogy már **tegnap** elindu**lok**,
originally so was that already yesterday away-start-pres.1sg
és akkor **holnapra** már ott is le**ttem** volna.
and then by.tomorrow already there too be-past-1sg expl-cond
'Originally, I was supposed to leave yesterday, and then I would have been there by tomorrow.'

(c) Péter azt mondta / rájött,
P. it-acc say-past.3sg / realize-past.3sg
[hogy Mari alszik/aludt].
that M. sleep-pres(3sg)/sleep-past(3sg)
'Peter said / realized that Mary was/had been asleep.'

  - Temporal adverbials do not formally agree with T; this is shown by the fact that they cannot locally license any T-switch in ellipsis under agreement (13a), the way subjects can do so with respect to Agr-marking (13b) (Bartos 2000):

(13) (a) Péter tegnap érkezett, Pál pedig ma [~~érkezett~~] / *[~~érkezik~~].
P. yesterday arrive-past(3sg) Paul and today arrive-past(3sg) / arrive-pres(3sg)
'Peter arrived yesterday, and Paul [arrived / *is arriving] today.'

(b) Tegnap Péter érkezett meg, ma pedig én [~~érkeztem~~].
yesterday P. arrive-past(3sg) pv today and I arrive-past-1sg
'Yesterday Peter arrived, while today I did.'

- As É. Kiss (2004; 2005) argued, what is regarded as tense-morphology in present-day Hungarian was originally an aspect marker (perfective), and has been reanalysed in the course of language change—but an alternative interpretation of historical facts, more consistent with the above two observations, is that tense-morphology is still rather like aspect marking: T-anterior vs. T-concurrent.
- Regarding "past tense" morphology as perfectivity marking is also in line with Eszes' (2004) analysis: he follows Condoravdi's (2001) ideas in analysing 'past' as 'perfective' in Hungarian.

In sum: tense in Hungarian is not well-suited for marking *excl*(x), hence this language resorts to the use of a particular mood to express *excl*(w) instead of T.

6. Conclusion

I have argued that conditional mood morphology in Hungarian is a reflex (spellout) of the feature *excl*(w), the semantic interpretation of which is counterfactuality. In certain cases (such as non-overtly-embedded wishes, or the antecedents of CF conditionals) this feature is lexically specified on M, with the direct interpretive aim of counterfactuality, while in others M has this feature unvalued, and inherits its value from the category Mod in a standard AGREE relation. The relevant feature of Mod, in turn, is possibly a derivative of its Kratzerian modality feature configuration, albeit this idea has not been pursued here. Finally, I have shown that, unlike what is found in many other languages, Hungarian cannot use tense marking as the exponent of *excl*(w), because its "tense" is relative, rather than deictic, i.e., it is rather aspect-like, so it is incapable of expressing exclusion (even *excl*(t), let alone *excl*(w)). Therefore this language makes use of mood morphology to encode CF, in the particular form of conditional mood.

Appendix

There is an alternative morphosyntactic account of *excl*-marking in Hungarian, which is more consistent with Iatridou's (2000) views. Following

É. Kiss (2004; 2005), and Iatridou (2000), one might speculate that conditional mood is morphologically complex in Hungarian: the affix *-nA* is segmentable into *-n-* and *-A*, the former being the mood affix proper, and the latter being none else than the now "extinct", archaic "narrative" past affix (cf. the forms *akara, esék, akará* in (6a,b) above):

(14) olvas-n-a cf. olvas-a
read-cond-past read-past

This way we would have an explanation for why "conditional mood" is used for encoding *excl(w)*. But there is precious little independent motivation for synchronically identifying the final *-A* as past tense, since this past tense is no longer in use in modern Hungarian, except, possibly, for this particular case in question. So there is practically no evidence available to the child at the time of language acquisition for having to segment the conditional marker as 'cond + past'. (NB. Such a segmentation account would have repercussions for the analysis of the syntax of inflection in Bartos (2003), as well, since this would then constitute motivation for assuming a $[T_{past}\ [M_{cond}]]$ hierarchy, rather than the other way round, making the case of dummy *vol-* insertion even more quirky—this case would then fall under some morphological constraint banning double tense-marking on a single V, most probably.) For this reason, I see no point in pursuing this idea any further.

References

Antal, László 1961. Gondolatok a magyar igéről [Thoughts on the Hungarian verb]. In: Magyar Nyelv 57:273–79.

Aoun, Joseph – Yen-hui Audrey Li 1993. Syntax of scope. MIT Press, Cambridge MA.

Baker, Mark 1985. The Mirror Principle and morphosyntactic explanation. In: Linguistic Inquiry 16:373–415.

Bartos, Huba 1999. Morfoszintaxis és interpretáció: a magyar inflexiós jelenségek szintaktikai háttere [Morphosyntax and interpretation: The syntactic background of Hungarian inflectional phenomena]. Doctoral dissertation, Eötvös Loránd University (ELTE), Budapest.

Bartos, Huba 2000. VP-ellipsis and verbal inflection in Hungarian. In: Acta Linguistica Hungarica 47:3–23.

Bartos, Huba 2003. On-line morphology: The morphosyntax of Hungarian verbal inflection. In: Uwe Jughanns – Luka Szucsich (eds): Syntactic structures and morphological information, 25–56. Mouton de Gruyter, Berlin & New York.

Cho, Eun 1997. Counterfactuals in Korean and Japanese: Interaction between verbal morphology and interpretation. Manuscript. Cornell University.

Chomsky, Noam 2001. Derivation by phase. In: Michael Kenstowicz (ed.): Ken Hale: A life in language, 1–52. MIT Press, Cambridge MA.

Cinque, Guglielmo 1999. Adverbs and functional heads: A cross-lingustic perspective. Oxford University Press, New York & Oxford.

Condoravdi, Cleo 2001. Temporal interpretation of modals. In: David Beaver – Stefan Kaufmann – Brady Clark – Luis Castillas (eds): Stanford papers on semantics, 1–30. CSLI Publications, Stanford.

Drubig, Hans Bernhard 2001. On the syntactic form of epistemic modality. Manuscript. University of Tübingen. (http://www.sfb441.uni-tuebingen.de/b2/papers/DrubigModality.pdf)

É. Kiss, Katalin 2004. A nézőponti aspektus jelölésétől a szituációs aspektus jelöléséig [From marking viewpoint aspect to marking situational aspect]. Paper presented at the Seventh International Congress of Hungarian Linguistics, August 2004, Budapest.

É. Kiss, Katalin 2005. Az ómagyar igeidőrendszer morfoszintaxisáról [On the morphosyntax of the tense system of old Hungarian]. Paper presented in the Research Institute for Linguistics, Hungarian Academy of Sciences, 10 February 2005, Budapest.

Eszes, Boldizsár 2004. Világidők és lehetőségek: új fejlemények a modális eseményszemantikában [World times and possibilities: New developments in modal event semantics]. In: Balázs Sinkovics (ed.): LingDok 4. Nyelvész doktoranduszok dolgozatai [LingDok 4. Papers by students of linguistics], 23–45. JATEPress, Szeged.

Farkas, Donka F. 1992. Mood choice in complement clauses. In: István Kenesei – Csaba Pléh (eds): Approaches to Hungarian 4 (The structure of Hungarian), 207–25. JATE, Szeged.

Iatridou, Sabine 2000. The grammatical ingredients of counterfactuality. In: Linguistic Inquiry 31 : 231–70.

Kálmán, László (ed.) 2001. Magyar leíró nyelvtan. Mondattan I [Hungarian descriptive grammar. Syntax 1] (Segédkönyvek a nyelvészet tanulmányozásához VI [Supplementary textbooks on linguistics 6]). Tinta Kiadó, Budapest.

Karttunen, Lauri 1972. Possible and must. In: J. P. Kimball (ed.): Syntax and semantics, vol. 1, 1 20. Academic Press, New York.

Kiefer, Ferenc 1981. What is possible in Hungarian? In: Acta Linguistica Hungarica 31 : 147–85.

Kiefer, Ferenc 1985. A *-hat/-het* képző jelentéstanához: az episztemikus *-hat/-het* [On the semantics of the modal suffix *-hat/-het*: Epistemic *-hat/-het*]. In: Általános Nyelvészeti Tanulmányok 16 : 131–53.

Kiefer, Ferenc 2005. Lehetőség és szükségszerűség. Tanulmányok a nyelvi modalitás köréből [Possibility and necessity. Studies in the domain of linguistic modality]. Tinta Könyvkiadó, Budapest.

Kratzer, Angelika 1981. The notional category of modality. In: Hannes Rieser – Hans-Jürgen Eikmeyer (eds): Words, worlds, and contexts, 38–74. Walter de Gruyter, Berlin.

Kratzer, Angelika 1991. Modality. In: Arnim von Stechow - Dieter Wunderlich (eds): Semantik/semantics: An international handbook of contemporary research, 639–50. Walter de Gruyter, Berlin.

Rákosi, György 2001. Mód és modalitás a magyarban [Mood and modality in Hungarian]. Paper presented at the Fourth Linguistics Conference for Doctoral Students (LingDok 4), Szeged.

Rebrus, Péter 2000. Morfofonológiai jelenségek [Morphophonological phenomena]. In: Ferenc Kiefer (ed.): Strukturális magyar nyelvtan 3. Morfológia [A structural grammar of Hungarian, Vol. 3: Morphology], 763–947. Akadémiai Kiadó, Budapest.

Tóth, Enikő 2006. Az alárendelt mondatbeli módválasztásról [On mood choice in subordinate clauses]. In: Balázs Sinkovics (ed.): LingDok 5. Nyelvész-doktoranduszok dolgozatai [Papers by doctoral students of linguistics], 181–200. JATEPress, Szeged.

Westmoreland, Robert R. 1995. Epistemic must as evidential. In: Paul Dekker – Martin Stokhof (eds): Proceedings of the 10th Amsterdam Colloquium, Part III, 683–702. ILLC, University of Amsterdam, Amsterdam.

Acta Linguistica Hungarica, Vol. 53 (3), pp. 249–289 (2006)
DOI: 10.1556/ALing.53.2006.3.2

A TYPOLOGY OF HUNGARIAN TIME ADVERBS*

ANIKÓ CSIRMAZ

Linguistics Program
Carleton College
One North College St
Northfield, MN 55057
USA
acsirmaz@carleton.edu

Abstract: Hungarian has a number of apparently synonymous time adverbs that can measure the duration of time intervals. The paper explores these adverbs in some detail, and argues that contrary to appearances, none of them are freely interchangeable. The starting point is a discussion of the property of homogeneity that time adverbs are sensitive to. The paper argues for a specific treatment of homogeneity and a preliminary adverb definition based on that treatment. It is proposed that some, but not all, Hungarian time adverbs share the default definition. The diverging adverbs may (a) contain a covert frequency predicate or (b) not measure the duration of the time interval directly, but by determining an endpoint of the interval. Hungarian time adverbs also differ in the range of time intervals they can measure; some, but not all adverbs can measure all available time intervals including the event, iterative, habitual and reference time. This variability in time adverb modification is arbitrary and needs to be explicitly determined for each adverb. Apart from discerning the interpretation of Hungarian time adverbs, the conclusions have a more general impact. On the one hand, apparently homogeneous adverbs can have disparate definitions. On the other, it is necessary to permit explicit, arbitrary constraints on adverbial modification. It is also argued that time adverbs can impose non-local restrictions on the eventuality modified, strengthening the need for a powerful theory of adverbial modification.

Keywords: aspect, duration, divisibility, adverbial modification, time interval

* For comments and discussions, I am grateful to Kai von Fintel, Danny Fox, Irene Heim, Sabine Iatridou, David Pesetsky, Katalin É. Kiss, Chris Piñón and the audience of the 30th Penn Linguistics Colloqium, the GLOW 29 Workshop on Adjuncts and Modifiers, the 16th Colloquium of Generative Grammar and of the workshop The Hungarian Language: Past and Present. The research reported here was partially funded by the OTKA grants TS 40705 and 49873. All errors are mine.

The present paper attempts to give a detailed characterization of Hungarian durative time adverbs. It is argued that the time adverbs which appear to be synonymous are not interchangeable; they encode a number of differences among one another. It is also pointed out that some Hungarian time adverbs impose unexpected and crosslinguistically marked restrictions on their arguments. More generally, the paper argues for varying treatments of homogeneity in eventuality descriptions and also argues for widening the range of time intervals that can be measured by time adverbs.

The paper is organized as follows. Section **1** introduces two classes of Hungarian and English time adverbs which are discussed in this paper. Section **2** elaborates on the properties of eventuality predicates and other predicates that the time adverb modification is sensitive to. Section **3** establishes a definition of English time adverbs, which also serves as the preliminary definition of Hungarian time adverbs. Sections **4** and **5** modify the preliminary definitions to describe the differences among Hungarian time adverbs and section **6** concludes the paper.

1. Introduction

The main focus of the paper is the group of adverbs that measure the duration of an eventuality, a category that includes both states and events, the latter a dynamic eventuality (Bach 1986). As often noted, these adverbs are sensitive to aspectual properties of the eventuality description. The type of eventuality description whose time the adverbs can measure is restricted; they either measure the time of an atelic eventuality description, or that of a telic one, as shown in (1).[1]

Before discussing the distribution of Hungarian time adverbs, let us briefly address the distinction between telic and atelic eventuality descriptions. One and the same event can be characterized in radically different ways: the eventuality of János running, for instance, can be described as in (1a) or as in (1b). A discussion of adverbial modification thus needs to appeal to properties of eventuality descriptions and not to those of events.

[1] The discussion is restricted to time adverbs that measure duration; punctual time adverbs are not addressed. Thus for the ease of discussion (and since the term *durative adverb* is sometimes restricted to adverbs measuring the time of atelic eventuality descriptions) I adopt the term *time adverb* to all adverbs measuring duration and apply it accordingly.

(1) (a) János ran ATELIC
(a′) János ran *for half an hour*
(b) János ran to the store TELIC
(b′) János ran to the store *in half an hour*

The two descriptions have different properties, which are discussed in more detail in section **2**. Intuitively, the atelic eventuality description in (1a) is homogeneous. As such, it can be applied, for instance, not only to a given eventuality, but also to a part or continuation of that eventuality. If János continues running after an hour and a half, then the longer running eventuality can still be described as *János ran.* Similarly, the eventuality description of *János ran* also holds during all parts of the 30-minute interval described. Telic eventuality descriptions, such as (1b), behave differently. The eventuality description *János ran to the store* cannot be applied to all proper parts of the running event. Similarly, if János runs further, then the telic eventuality description cannot apply to this larger eventuality.

This homogeneity difference between telic and atelic eventuality descriptions is shown by a number of diagnostics (e.g., Smith 1991; Rothstein 2004). One of the most often cited diagnostics relies on temporal modification of eventualities. The duration of the event argument of an atelic eventuality description can be measured by a *for*-adverb (2a). That of an argument of a telic description, in contrast, is modified by an *in*-adverb (3a).

In Hungarian, four different durative adverbs can modify an atelic eventuality description (2b d). I argue below that contrary to the initial impressions, these adverbs are not synonymous. In fact, the distribution or the interpretation of all of these adverbs is different, and they can impose different restrictions on the time intervals they measure.

(2) Adverbs measuring the time of an atelic eventuality description
(a) János ran / *ran to the store *for an hour and a half*
(b) János *másfél* *órán* *át* futott / *el futott a boltba
J-nom one.and.half hour-on across ran away ran the store-to
'János ran / *ran to the store for an hour and a half'
(c) János *másfél* *óráig* futott / *el futott a boltba
J-nom one.and.half hour-until ran away ran the store-to
'János ran / *ran to the store for an hour and a half'

(d) János *másfél* *órán* *keresztül* futott / *el futott a boltba
J-nom one.and.half hour-on across ran away ran the store-to
'János ran / *ran to the store for an hour and a half'

(e) János *másfél* *órát* futott / *el futott a boltba
J-nom one.and.half hour-acc ran / away ran the store-to
'János ran / *ran to the store for an hour and a half'

For telic eventuality descriptions, two types of adverbs modify duration in Hungarian. Similarly to the time adverbs in (2), I argue below that these adverbs are not synonymous but affect time intervals differently.

(3) Adverbs measuring the time of a telic eventuality description

(a) János ran to the store / *ran *in an hour and a half*

(b) János *másfél* *óra* *alatt* el futott a boltba / *futott
J-nom one.and.half hour under away ran the store-to ran
'János ran to the store / *ran in an hour and a half'

(c) János *másfél* *órán* *belül* el futott a boltba / *futott
J-nom one.and.half hour-on inside away ran the store-to ran
'János ran to the store / *ran in an hour and a half'

Before turning to a detailed discussion of Hungarian time adverbs, let us discuss the characterization of (a)telicity and the semantics of time adverbs below. Section **2** explores diverse views of homogeneity and section **3** discusses the semantics of time adverbs in general.

2. Properties of eventuality descriptions

In order to characterize the difference between telic and atelic eventuality descriptions reliably, it is necessary to identify certain properties of eventuality descriptions. This section argues that the homogeneity of eventuality descriptions is best characterized in terms of divisibility rather than cumulativity. It is pointed out that the standard definition of divisibility or subinterval property runs into problems, and an alternative definition is adopted. It is also shown that the relevant notion of homogeneity cannot apply only to eventuality descriptions. It must be applicable to other predicates of times, such as those that take the reference time or the perfect time span as an argument.

2.1. Divisibility or cumulativity?

As noted in the preceding section, atelic eventuality descriptions are homogeneous. Homogeneity is usually described in terms of divisibility and/or cumulativity, both of which are properties of predicates (Smith 1991; Rothstein 2004; Krifka 1998, among others). Divisible predicates hold for a given argument and all of its parts. Cumulative predicates, in turn, apply not only to atomic arguments, but also to their union.

(4) (a) A predicate P is divisible iff whenever $P(x)$, then $\forall y \subseteq x \rightarrow P(y)$
(b) A predicate P is cumulative iff whenever $P(x)$ and $P(y)$, then $P(x \oplus y)$[2]

In order to apply the definitions of divisibility or cumulativity to eventuality descriptions, it is necessary to introduce specific assumptions. I assume that eventuality descriptions take, among others, a time interval argument t, the event time.[3] The predicate of times applying to the event time is the event time predicate. Divisibility and cumulativity, as defined in (4), apply to the predicate of times (P) and the time interval argument of that predicate (t), as illustrated below.

(5) (a) *János run* is divisible
(János run)$(t) \rightarrow \forall t' \subset t \rightarrow$ (János run)(t')
(b) *János run* is cumulative
(János run)(t) & (János run)$(t') \rightarrow$ (János run)$(t \oplus t')$[4]

Since both divisibility and cumulativity hold only of atelic eventuality descriptions but not of telic ones, either property appears to be sufficient to distinguish the two types of predicates. With atelic eventuality descriptions, divisibility and/or cumulativity holds for the event time predicate (P) and the event time (t). If the eventuality description is telic, then the event time predicate is neither divisible nor cumulative.[5]

[2] I assume that eventualities, similarly to locations and individuals, can be temporally and spatially discontinuous.

[3] In this paper, I am agnostic about whether the predicates have an event argument, or even whether events exist. Appealing to time intervals rather than events in determining telicity allows for a more general treatment. Homogeneity can be straightforwardly extended to predicates that apply to times other than the event time (e.g., the reference time), as discussed below.

[4] Time intervals—like eventualities—can be discontinuous.

[5] The view of telicity as the property of the event time predicate (in contrast with (im)perfectivity, as discussed below) assumes a two-component theory of

As they stand, however, neither property identifies the range of atelic and telic eventuality descriptions properly. First, let us consider granularity, a shortcoming of the property of divisibility.

Given the predicate of times of an atelic eventuality description, that predicate does not necessarily apply to all the subintervals of that argument (discussed in Hinrichs 1985; Rothstein 2004; Bertinetto 2001; among others). Consider the atelic examples in (6).

(6) (a) János ran (for ten minutes)

(b) János futott
J-nom ran
'János ran'

(c) János was sick (for two days)

(d) János beteg volt
J-nom sick was
'János was sick'

A state, as in (6c, d), holds for all the subintervals of the time argument, here a two-day-long time interval. For an activity such as running (6a, b), in contrast, this is not the case. Activities show the granularity effect: the time interval argument contains atomic time intervals for which the predicate of times—in this case *János run*—is not true. The predicate fails to hold, among others, for the time interval during which he only lifts his right heel off the ground.[6] Since divisibility, as defined above, requires the predicate in question to hold of all parts of the argument, a number of atelic eventuality descriptions—specifically, all activities—fail to qualify as divisible.

---

aspect, as in Smith (1991), Olsen (1997), Bertinetto (2001), and others. In this theory, the property of telicity ('situation aspect') is crucially distinct from that of (im)perfectivity (the 'viewpoint aspect'). The difference is encoded here as homogeneity applying to distinct time intervals: the event time for telicity, and the reference time for (im)perfectivity (cf. section **2.2**).

[6] The lack of the granularity effect has been suggested as the property (or one of the properties) distinguishing states and dynamic, non-stative divisible eventuality descriptions (for instance, Vendler 1967; Dowty 1979; Comrie 1976; Smith 1991; and Bertinetto 2001). As argued by Csirmaz (to appear), non-granularity does not hold of states only, but can also be true of other predicates of times, including reference time intervals discussed in section **2.2.** Homogeneity without granularity thus cannot identify stative descriptions. Rather, states can be identified by either (a) restricting strict, non-atomic homogeneity to event predicates, or (b) by appealing to a different property such as inertia (lack of dynamicity), a property that holds only of states.

As noted above, telic eventuality descriptions are not cumulative. It may suffice then to adopt cumulativity, repeated below, as the relevant property distinguishing telic and atelic eventuality descriptions.

(7) A predicate P is cumulative iff whenever $P(x)$ and $P(y)$, then $P(x \oplus y)$

Cumulativity does not encounter the problem of granularity. It identifies *János run* as cumulative and therefore atelic, in spite of the existence of atomic time intervals where the predicate *János run* does not hold. While it handles the distinction between atelic and telic eventuality descriptions successfully, cumulativity fails in predicting the range of adverbs modifying other time intervals. A time interval other than the event time and its interaction with time adverbs, cumulativity and divisibility is addressed in the following section.

2.2. Reference time and predicates

The preceding discussion was concerned with properties of the event time predicate and adverbial modification. It is not only the event time that can be measured by time adverbs, though. In (8), for instance, it is not the event time but the perfect time span that is modified; there is a two-year-long interval, during which János lived in Spain, which extends in the past from the speech time backwards.

(8) János has lived in Spain for two years

Apart from the event and perfect times, other time intervals can also be measured by time adverbs.

2.2.1. Reference time

Csirmaz (2005; 2006; to appear) notes that the reference time can also be modified by a time adverb. The reference time—or topic time—is the time interval under discussion, which can be ordered in a number of different ways with respect to the event time and the time of utterance. As argued by Klein (1994), Iatridou et al. (2001), von Stechow (2002) and others, the relative ordering of the reference time and the event time yields the perfective or imperfective viewpoint of an eventuality description.[7]

[7] Demirdache and Uribe-Etxebarria (2000; 2004) argue for a related but essentially different view of time intervals. They assume that the possible orderings for time

(9) (a) János ran
(b) János was running

In the perfective (9a), the event time is properly contained within the reference time, hence the intuition that the sentence focuses on or asserts the complete event of running. In the imperfective (9b), in contrast, the reference time is a proper subinterval of the event time. An imperfective eventuality description thus focuses on a part of the eventuality rather than on the eventuality as a whole. The definitions, based on Iatridou et al. (2001) and von Fintel–Iatridou (1997), are given below.

(10) (a) $[\![\text{perfective}]\!] = \lambda P.\lambda t.\exists t'.[t' \subset t \,\&\, P(t')]$
(b) $[\![\text{imperfective}]\!] = \lambda P.\lambda t.\exists t'.[t \subset t' \,\&\, P(t')]$
(t: reference time; t': event time)

Similarly to the event time, the reference time can also be measured by a time adverb. In English, it is possible for both the event time and the reference time to be modified at the same time (as noted in de Swart 1998, for instance):[8]

(11) For half an hour, János was running the distance in ninety minutes
(but then he realized that he wouldn't be able to complete it in time)

The adverb *for half an hour* measures the reference time of the imperfective eventuality description, and *in ninety minutes* specifies the duration of the event time. Thus, as shown by the possibility of perfect and reference time modification, the treatment of time adverbs should not appeal to properties of the event time only (contrary to Moltmann 1991, among others).

---

intervals are consecutive ordering and inclusion. For the event time and their assertion time, this ordering yields either a perfect or an imperfective eventuality description. The system makes it impossible to represent and account for (non-perfect) perfective eventuality descriptions. Csirmaz (2006) argues that given the existence of non-perfect perfective eventuality descriptions, and that of languages that overtly mark such descriptions, it is more attractive to adopt the system outlined above. The topic and event times can be ordered by inclusion, and the perfect time is optionally present in the time structure of the description. The reference time and the time of utterance can be ordered by precedence or containment, as in past and present tense eventuality descriptions, respectively.

[8] Similar examples are ungrammatical in Hungarian, as discussed in section 5.

2.2.2. Homogeneity as divisibility

The modification of reference time helps to refine the criteria for determining predicate homogeneity. As pointed out earlier, the granularity of atelic event descriptions argues against determining homogeneity in terms of divisibility. If homogeneity is seen as cumulativity, then the problem of granularity and atomic time intervals does not arise. The possibility of reference time modification, however, suggests that it is divisibility that should be viewed as the relevant criterion of homogeneity.

It was noted by Bennett–Partee (1972) and Dowty (1979) among others that all negated eventuality descriptions can be modified by a *for*-adverb, including telic and perfective descriptions. *For*-adverb modification shows that negation yields a homogeneous eventuality description:

(12) (a) *For half an hour*, János didn't arrive
(a′) #*For half an hour*, János arrived
(b) János *fél óráig* nem érkezett meg
J-nom half hour-until not arrived perf
'For half an hour, János didn't arrive'
(b′) #János *fél óráig* meg érkezett
J-nom half hour-until perf arrived
'For half an hour, János arrived'

In addition, *for*-adverbs can also modify all eventuality descriptions with a decreasing argument or those with a constituent modified by *only*. Again, the telicity and perfectivity of the eventuality description is irrelevant. The eventuality descriptions below are telic and perfective, yet modification by a *for*-adverb or a Hungarian counterpart is grammatical.

(13) (a) For half an hour, fewer than three guests arrived
(b) Fél óráig kevesebb, mint három vendég érkezett meg
half hour-until fewer than three guest-nom arrived perf
'For half an hour, fewer than three guests arrived'

(14) (a) For four months, only János completed the course (the others didn't manage to do so)
(b) Négy hónapig csak János végezte el a tanfolyamot
four month-until only J-nom completed away the course-acc
'For four months, only János completed the course'

Csirmaz (2005; 2006; to appear) argues that in these cases the time adverbs modify the reference time, whereas the properties of the event time —including telicity and duration—remain unaffected. Modification by *for*-adverbs and their Hungarian equivalents is thus not a diagnostic of the homogeneity of eventuality descriptions. Rather, the time adverbs can ascertain the homogeneity of diverse predicates of time—that of the predicate applying to the event time, the reference time, or the perfect time span.

Let us assume that the semantics of *for*-adverbs is uniform, and that *for*-adverbs only modify predicates of times that are homogeneous in a certain uniform sense. Given this assumption, the preceding data enforce the view of homogeneity that appeals to divisibility rather than to cumulativity. Consider the eventuality description with a monotone decreasing argument, as in (13). The *for*-adverb measures the duration of the reference time rather than that of the event time, since the event time predicate is not homogeneous.

The predicate of times which applies to the reference time in (13) is clearly not cumulative. If two guests arrived during a time interval t and one guest arrived during time t', then for both t and t' it holds that fewer than three guests arrived during those intervals. During the union of t and t', however, it is exactly three guests that arrived—the predicate *fewer than three guests arrived* thus does not hold for $t \oplus t'$. Thus even though cumulativity fails to encounter the problem of granularity, it cannot identify the set of homogeneous predicates of times, which can be modified by a *for*-adverb.[9]

2.3. Two approaches to divisibility

2.3.1. Granularity and contextual restriction

Since cumulativity cannot determine homogeneity for time adverb modification, divisibility must be adopted instead. It is necessary then to avoid the granularity problem by altering the original definition of divisibility, repeated below.

[9] Divisibility as the property determining homogeneity does not account for the *for*-adverb modification licensed by *only*, illustrated in (14). To account for this, I introduce the notion of Strawson divisibility (building on Strawson entailment (von Fintel 1999)), which requires divisibility to hold only for those time intervals where the predicate of times is defined. Strawson divisibility is discussed in more detail in section **3.3.**

(15) A predicate P is divisible iff whenever $P(x)$, then $\forall y \subseteq x \rightarrow P(y)$

A number of authors (more recently Moltmann 1991; Bertinetto 2001; Rothstein 2004) suggested that the universal quantification over parts of arguments still holds, but is constrained by some contextual restriction. This restriction ensures that the predicate of times does not need to hold for all subintervals, but only for those which are not excluded by that restriction. Even though this appears to be a viable way to salvage divisibility as the relevant condition of homogeneity, a number of problems arise.

First, as noted by Hinrichs (1985), the nature of the contextual restriction is a highly pragmatic matter. The length of atomic time intervals for the predicate *János run* can be affected, among others, by the age or physical properties of *János*, the agent. If divisibility is treated as a semantic property, however, then these pragmatically affected restrictions cannot be incorporated.

In addition, even within the domain of semantics proper, granularity (the existence of atomic time intervals) leads to circularity: the contextual restriction approach must assume that whenever a divisible predicate applies to an argument, it must also apply to all parts of that argument to which the predicate could apply. As an illustration, consider the following example:

(16) (a) the statue stood on the square
(b) a szobor a téren állt
the statue-nom the square-on stood
'the statue stood on the square'

The stative description *the statue stand on the square* is divisible – an uncontroversial matter, since the event predicate is not granular. The eventuality description may contain a non-divisible time adverb, as in (17).[10] The resulting eventuality description becomes non-divisible, since the duration of the eventuality description is delimited.

(17) (a) the statue stood on the square for a hundred years
(b) a szobor *száz évig* állt a téren
the statue-nom hundred year-until stood the square-on
'the statue stood on the square *for a hundred years*'

[10] Specifically, the time adverb takes a time interval argument and the adverb is non-divisible with respect to that time interval.

Whenever the time adverb is divisible, as with the bare plural *centuries* or its Hungarian equivalent, the predicate is still homogeneous—that is, divisible—after temporal modification:

(18) (a) the statue stood on the square for centuries

(b) a szobor *évszázadokig* állt a téren
the statue-nom centuries-until stood the square-on
'the statue stood on the square for *centuries*'

The homogeneity of the resulting eventuality description cannot be shown by adverbial modification. Nevertheless, intuitively the homogeneity holds for the description, since the endpoint of the time during which the description holds is not specified. Assuming that homogeneity can always be equated with divisibility, the eventuality description in (18) must be divisible.

The time interval during which the divisible eventuality description holds is highly granular; while the description is homogeneous, homogeneity does not apply for all subintervals. The atomic subintervals are one hundred year long, since the atoms are those intervals to which the predicate *egy évszázadig* ('for a century') or *century* can apply.[11] Given this condition on atomic times, (18) is homogeneous, similarly to other iterative or habitual eventuality descriptions.

To account for examples such as (18), an approach that appeals to contextual restriction on universal quantification needs to assume that the atomic time intervals are those intervals for which the predicate can hold. In the present case, the atomic time intervals measure one hundred years, as noted above. This restriction of atomic time intervals is rather circular, since it determines those predicates as divisible which satisfy the following condition: the predicate applies to an argument and all parts of that argument to which in can apply.

The circularity of atomic time intervals and arguments is also found elsewhere; it is revealed by all predicates of time that have sufficiently long atomic intervals. The following examples are all homogeneous and can be modified by *for*-adverbs and certain Hungarian equivalents of these adverbs.[12] As before, the homogeneity of the predicate is ensured

[11] Bare count nouns are cumulative in Hungarian, thus a numeral or determiner is required to denote a time interval that is 100 years long.

[12] The restriction on Hungarian equivalents of *for*-adverbs in modifying habitual and iterative eventuality descriptions is discussed in section 4 in more detail.

only if the atomic intervals are defined as intervals which are possible arguments of the predicate.

(19) (a) János built churches

(b) János templomokat épített
J-nom churches built
'János built churches' (possible interpretation: 'built complete churches')

(20) (a) János wrote novels

(b) János regényeket írt
J-nom novels-acc wrote
'János wrote novels' (possible interpretation: 'wrote complete novels')

Thus both the pragmatic restriction and circularity present a problem for the approach that assumes a contextual restriction on universal quantification.

2.3.2. A modified view of divisibility

In order to avoid the problem created by granularity and the non-feasibility of contextual restriction on atomic arguments, Hinrichs (1985) and von Fintel (1997) propose a more complex definition of divisibility.

(21) A predicate P is divisible iff whenever $P(x)$ for an argument x, then for all $y \subset x, \exists z[y \subseteq z \subset x \,\&\, P(z)]$
(all proper parts of x must be parts of P-arguments) (based on Hinrichs 1985)

(22) A predicate P is divisible iff whenever $P(x)$ for an argument x, then $x = \oplus_{\mathrm{NT}}\{y : P(y)\}$
(x is the (non-trivial) sum of a set of P-arguments) (von Fintel 1997)

Both definitions ensure that a time interval which serves as the argument of a divisible predicate of times has at least two disjoint subintervals that are also arguments of the predicate. This definition of divisibility avoids the problem of atomic predicates noted above, since the predicate does not need to apply to all subintervals. In addition, unlike cumulativity, it permits identifying the reference time of predicates with a decreasing argument as homogeneous.

Adopting this view of divisibility, the correlations between predicates of times and adverbial modification can thus be noted as in (23).

(23) (a) A divisible predicate of times can be modified by a *for*-adverb
(b) A non-divisible predicate of times can be modified by an *in*-adverb

Note that no claims have been made about the distribution of the Hungarian equivalents of these adverbs. As emphasized earlier, I will argue that the distribution of Hungarian time adverbs differs among the various time adverbs. The generalizations concerning the distribution of English time adverbs and the definitions of these adverbs, discussed below, serve as a starting point for the differences among Hungarian time adverbs.

3. Time adverbs and time intervals

In the preceding section I argued that the homogeneity of atelic eventuality descriptions is best described as divisibility in the sense of Hinrichs (1985) and von Fintel (1997). Divisibility extends not only to event time predicates (distinguishing telic and atelic eventuality descriptions), but also to perfect and reference time predicates. *For*-adverbs can thus measure event times, reference times and perfect times, assuming that the predicate applying to the time interval is divisible. *In*-adverbs, in contrast, measure a time interval if the predicate of times is non-divisible.

Before turning to Hungarian time adverbs, let us determine a definition for the English *for* and *in*-adverbs. These will be adopted as the preliminary definitions for Hungarian time adverbs.

3.1. A first approach

Of the two time adverb classes, let us discuss *for*-adverbs and their equivalents—henceforth A-adverbs—first. As before, I assume the existence of time intervals (t) and predicates of time (P). Based on the preceding discussion, the approaches that assume the standard definition of divisibility (4a) or divisibility constrained by contextual restriction all encounter problems with granularity. Thus the accounts of Zucchi (1991), Moltmann (1991) and others, which incorporate this notion of divisibility in the denotation of a *for*-adverb, need to be modified.

An alternative definition of A-adverbs, based on (the divisibility definition of) Hinrichs (1985) and von Fintel (1997), is given below. The adverb takes a predicate of times and a time interval argument. The

predicate of times argument must be divisible, ensuring that only divisible predicates of time can be modified by A-adverbs.[13]

(24) for twenty minutes = $\lambda P.\lambda t.[\forall t' \subset t[\exists t''[t' \subseteq t'' \subset t \,\&\, P(t'')]] \,\&\, |t| = 20\,\text{minutes}]$

Even though this definition ensures that the adverb cannot modify non-divisible predicates of times, at first blush it seems to run afoul on iterative and habitual eventuality descriptions. Both of the latter can be modified by an A-adverb:[14]

(25) (a) János ran for three years (HABITUAL)

(b) János három évig futott
J-nom three year-until ran
'János ran for three years'

[13] The adverb also has a measure argument which specifies the length of the time interval. For simplicity, I treat this measure argument as part of the adverb in this paper. It is worth noting, however, that the measure arguments have a maximality implicature which can be cancelled:

(i) János fél óráig fel mosott
J-nom half hour-until up washed
'János washed the floor for half an hour'

(ii) Sőt, volt az egy óra is
even was that one hour too
'It lasted an hour, even'

The implicature cannot be cancelled if the adverb is in immediately preverbal position.

(iii) János fél óráig mosott fel
J-nom half hour-until washed up
'János washed the floor for half an hour'

(iv) #Sőt, volt az egy óra is
even was that one hour too
'It lasted an hour, even'

The effect of preverbal position on the cancellability of maximality implicatures is not unique to these adverbs, but holds for other constituents that introduce an implicature as well. See É. Kiss (in press) for a recent discussion.

[14] As discussed in section **4**, distinct types of Hungarian A-adverbs differ in whether they can modify an iterative or habitual eventuality description.

(26) (a) János knocked for ten minutes (ITERATIVE)

(b) János tíz percig kopogott
J-nom ten minute-for knocked
'Janos knocked for ten minutes'

Iterative and habitual eventuality descriptions can contain gaps where the event predicate does not hold. János does not need to run continuously during the three years in (25). Similarly, there can be times during the ten-minute interval in (26) when he is not knocking. The predicates of times applying to the three-year and ten-minute interval must then be divisible and still allow gaps.[15]

The notion of divisibility proposed earlier and the definition of A-adverbs above require that all subintervals of the time argument t—including gaps—be contained in some interval t' for which the predicate of times holds. This requirement fails for instantaneous eventualities such as the event *János knock*, which only holds for atomic, momentary time intervals. Gaps between running eventualities, as in (25), are also different from the atoms that the definition of divisibility handles successfully; it is not necessarily the case that the predicate of times *János run* extends over the three-year-long interval.

An alternative definition of A-adverbs explored in the following section readily accounts for the existence of gaps. It is argued that the original definition is more attractive, and that the problem of gaps can be resolved with this definition as well.

3.2. Gaps and divisibility

The definition of A-adverbs, as assumed above, does not seem to permit modification of a time interval containing gaps. The incompatibility of these adverbs and gaps is predicted because, given the definition of divisibility in (27), the maximal time intervals of habitual and iterative eventuality descriptions are non-divisible with respect to the event time predicate.

[15] Gaps differ from atomic time intervals, which give rise to the granularity effect. For gaps time intervals, there is a time argument of the predicate of times that contains the gap. For gaps, no such interval needs to exist; the gap is a time interval which is both preceded and followed by other time intervals that serve as arguments of the predicate.

(27) A predicate P is divisible iff whenever $P(x)$ for an argument x, then for all $y \subset x, \exists z[y \subseteq z \subset x \,\&\, P(z)]$
(all proper parts of x must be parts of P-arguments) (based on Hinrichs 1985)

Contrary to what is predicted, both habitual and iterative eventuality descriptions permit modification by A-adverbs, as noted above and repeated below for the habitual (28) and the iterative (29).

(28) (a) János ran for three years
(b) ános három évig futott
J-nom three year-until ran
'János ran for three years'

(29) (a) The lamp blinked for ten minutes
(b) A lámpa tíz percig pislogott
the lamp-nom ten minute-until blinked
'The lamp blinked for ten minutes'

If the predicate of times argument of A-adverbs must be divisible, then the definition of divisibility needs to be revised. Divisibility must allow the time intervals modified to contain not only atomic time intervals but also gaps.

Piñón (1999), after pointing out these problems, suggests that *for*-adverbs neither measure the duration of some time interval nor involve quantification over subintervals. In order to account for the possible presence of gaps, he suggests that A-adverbs take a (possibly implicit) frequency predicate argument, which specifies the frequency of appropriate eventualities within the time interval in question. The frequency predicate R takes an eventuality, a time interval and an eventuality type as arguments. Eventualities of type P are repeated throughout the time interval t with the frequency specified (the relation of repetition explicitly specified by Piñón 1999).

(30) for twenty minutes $= \lambda R \lambda P \lambda e[\exists t[[20 \text{ minute}](t) \,\&\, R(e, t, P)]]$[16]

If there is no overt frequency predicate, the A-adverb can be interpreted differently. In that case it is possible that there is an event with a runtime that is coextensive with the time interval t, and for which the eventuality

[16] Piñón (1999) also lists the measure phrase as an argument of the *for*-adverb.

predicate is true. The event is not necessarily iterated or repeated, and thus the adverb can measure the duration of a single event.[17]

The adoption of a frequency predicate solves the problem of gaps. The problem of granularity is resolved by the assumptions concerning the alternative form of the adverb, which can appear in absence of an overt frequency adverb. In the latter case the eventuality predicate is not required to be true at all subintervals of the time interval modified, so the issue of atomic time intervals does not arise.

It was noted in section **2** that *for*-adverbs can modify not only event times, but also other time intervals. If Piñón's proposal is adopted without modifications, then *for*-adverbs are restricted to event time modification only. The definition can be modified such that the adverbs take not an eventuality, but a time interval argument. With the A-adverb denotation with overt frequency predicates, the frequency predicate applies to the time interval t, and P—a predicate of times—holds throughout t.

(31) for twenty minutes = $\lambda R \lambda P \lambda t[\exists t'[[20\ \text{minute}](t') \,\&\, R(t, t', P) \,\&\, P(t)]]$

Similarly, the alternative entry of the adverb can also be rephrased and refer to time intervals rather than events or event times.

While the suggested modification resolves the issue of restricted time modification, there are still two entries of A-adverbs that are necessary to account for the readings of the adverb. With a different view of the status of gaps, it may be possible to maintain a unique definition of A-adverbs.[18]

I suggest that the definition of A-adverbs based on Hinrichs (1985) and von Fintel (1997) is, in fact, an adequate definition that handles both gaps and atomic intervals. The definition of A-adverbs in question is repeated below from (24).

(32) for twenty minutes = $\lambda P.\lambda t.[\forall t' \subset t[\exists t''[t' \subseteq t'' \subset t \,\&\, P(t'')]] \,\&\, |t| = 20\,\text{minutes}]$

[17] The alternative denotation (with Q a measure predicate) from Piñón (1999) is given as follows:

$$\begin{aligned}&\text{for} = \lambda Q \lambda P \lambda e[\exists t[Q(t) \,\&\, \text{Rep}(e, t, P) \,\&\, \Diamond \exists e'[\tau(e') = \\ &= t \,\&\, P(e') \rightarrow \forall t'[\text{Pause}(t', t, e, P) \rightarrow \\ &\text{Interrupt}(t', t, e, P)] \,\&\, \neg \Diamond \exists e'[\tau(e') = t \,\&\, P(e')] \rightarrow \\ &\neg \text{Con}(e) \,\&\, \forall t'[t' \subseteq t \,\&\, S(t') \,\&\, \exists e'[e' \subseteq e \,\&\, (e') \subseteq t' \,\&\, P(e')]]\end{aligned}$$

[18] In section **4**, it is argued that the first entry of the *for*-adverb must be adopted (in a modified form) for some Hungarian A-adverbs. Thus while the "default" A-adverb definition is different, Pinón's definition must still be adopted in some cases.

Contrary to initial appearances, gaps do not present a problem for this definition; iterative as well as habitual eventuality descriptions qualify as divisible. They are divisible since the definition applies to the habitual and iterative predicates of times rather than to the event time predicates themselves. That is, the habitually or iteratively recurring event does not need to be expressed as a divisible event predicate. Rather, iterative and habitual eventuality predicates can hold of time intervals even when the event that recurs habitually or iteratively is not true.

(33) (a) [HAB (János run)](t)
(b) [ITER (lamp blink)](t)

Divorcing the iterative/habitual predicate from the event time predicate (which describes the iteratively/habitually recurring event) successfully accounts for the apparent problem of gaps. This view makes it necessary to establish not only event time and reference time, but also a habitual time and iterative time. In addition to these times, there are also predicates of times applying to these time intervals.

3.3. Downward entailing quantifiers, *only* and adverbs

Downward entailing quantifiers, mentioned in section **2.2**, also support adopting the definition based on Hinrichs (1985) and von Fintel (1997) and disfavors a modification of the treatment of Piñón (1999). Recall that downward entailing quantifiers permit A-adverb modification even if in absence of these quantifiers, A-adverbs are marked. Licensing is illustrated below, repeated from (13).

(34) (a) For half an hour, #(fewer than) three guests arrived
(b) Fél óráig #(kevesebb, mint) három vendég érkezett meg
half hour-until fewer than three guest-nom arrived perf
'For half an hour, fewer than three guests arrived'

It was observed above that divisibility, as defined based on Hinrichs (1985) and von Fintel (1997), handles these facts straightforwardly. In addition, a modification of divisibility also extends to A-adverb licensing by *only*, as in the example repeated from (14).

(35) (a) For four months, only János completed the course (the others didn't manage to do so)

(b) Négy hónapig csak János végezte el a tanfolyamot
four month-until only J-nom completed away the course-acc
'For four months, only János completed the course'

The reference time predicate, when modified by *only*, holds for only those subintervals that contain the event time. The predicate is not divisible according to the definition adopted; it is not true that all subintervals are contained in a proper subinterval for which the predicate holds. The definition of divisibility must thus be modified to ensure divisibility of this predicate.

A successful treatment of the A-adverb modification of reference time predicates with *only* requires several ingredients. Among others, (a) the introduction of the notion of Strawson divisibility, where the divisibility only needs to hold for a subset of the subintervals (for the subintervals where the predicate is interpreted)[19] and (b) a way of determining the duration of the reference time of predicates with *only*. An elaboration of such an account is outside of the scope of the present paper, but a possible treatment is described in Csirmaz (2005; to appear).

To summarize: a uniform treatment of A-adverbs is possible. The definition of A-adverbs must adopt divisibility based on the definitions of Hinrichs (1985) and von Fintel (1997). The resulting definition handles both atoms and gaps successfully. In addition, it extends to reference time predicates with downward entailing quantifiers and—if divisibility is modified and understood as Strawson divisibility—to predicates of times modified by *only*. The A-adverb can measure the duration of a number of time intervals—including the event time, iterative, habitual or reference time—if the predicate applying to these time intervals is divisible.

[19] The notion of Strawson divisibility builds on Strawson entailment (von Fintel 1999), where the entailment relation is similarly restricted. It is worth pointing out that Strawson divisibility also distinguishes an approach based on a modified notion of divisibility and that of Piñón (1999). While the divisibility account readily accounts for Strawson divisibility, it is not immediately clear how the latter approach accounts for these facts.

3.4. *In*-adverbs

In contrast with *for*-adverbs, *in*-adverbs (henceforth T-adverbs) modify telic, non-divisible eventuality descriptions. The condition on predicates of times is built into the definition, which differs from that of A-adverbs in requiring a non-divisible predicate of times argument.

(36) for twenty minutes = $\lambda P.\lambda t.[\forall t' \subset t[\exists t''[t' \subseteq t'' \subset t \,\&\, P(t'')]] \,\&\, |t| = 20\,\text{minutes}]$

(37) in twenty minutes = $\lambda P.\lambda t.[\neg[\forall t' \subset t[\exists t''[t' \subseteq t'' \subset t \,\&\, P(t'')]]] \,\&\, |t| = 20\,\text{minutes}]$

A telic eventuality description can be modified by a T-adverb because for some subintervals t' of the event time—namely, for those that include the left boundary of the event time—there is no proper subinterval t'' of t for which the predicate holds. For the event predicate *János go to the store*, for instance, the predicate holds of the event time t and also for those subintervals that include the endpoint of t. Crucially, the predicate only applies to those subintervals that include the endpoint. It follows then that those proper subintervals that contain the initial point of t cannot be parts of a proper subinterval of t for which *János go to the store* also holds.

For A-adverbs, it was proposed earlier that they can measure a number of time intervals and are not restricted to event time modification. The variability of modification can also hold of *in*-adverbs. It is possible to view *in*-adverbs as measuring either the event time or the reference time interval, as illustrated below.

(38) (a) János wrote the letter in an hour
(event time; the event lasted an hour)

(b) János arrived (with)in an hour
(reference time; the event occurred at some point within the hour-long interval)

The different time adverbs measured result in different interpretations of the T-adverb. If the event time is measured, then the event lasts as long as specified. If the adverb measures the reference time, then the event time is contained within the reference time.[20] A maximality implicature

[20] While *in*-adverbs are ambiguous in measuring either the event time or the reference time, *within*-adverbs can only measure the reference time. It must also be pointed out that English T-adverbs impose a restriction on the eventuality

account (where it is implicated, but not asserted, that the event lasted as long as specified by the adverb) can also account for the reference time modification cases. It will be shown in section **5**, however, that the account of Hungarian T-adverbs needs to appeal specifically to reference time modification by a T-adverb. Since the possibility is independently attested, it may be assumed that English T-adverbs can also modify the reference time.

In the following sections I take the previous definitions of A- and T-adverbs as starting point and note where their Hungarian equivalents diverge. I also assume that time adverbs can show flexibility in measuring time intervals, as shown for English A-adverbs and suggested for T-adverbs above.

4. A-adverbs in Hungarian

In the preceding section I argued for a specific definition of divisibility and definitions of English A- and T-adverbs, both based on Hinrichs (1985) and von Fintel (1997). For Hungarian time adverbs, it was pointed out that they are not synonymous, but each adverb shows a different distribution. The Hungarian A-adverbs, enumerated earlier, are italicized below.

(39) (a) János *másfél* *óráig* futott (*-ig* adverb)
J-nom one.and.half hour-until ran
'János ran for an hour and a half'

(b) János *másfél* *órán* *keresztül* futott (*keresztül* adverb)
J-nom one.and.half hour-on through ran
'János ran for an hour and a half'

(c) János *másfél* *órán* *át* futott (*át* adverb)
J-nom one.and.half hour-on across ran
'János ran for an hour and a half'

(d) János *másfél* *órát* futott (accusative adverb)
J-nom one.and.half hour-acc ran
'János ran for an hour and a half'

description: independently of the time interval modified, they require the event time predicate to be telic.

The apparently synonymous adverbs differ in various ways. I first consider the adverbs and the time intervals they can modify, suggesting that the range of times that an adverb may modify must be independently specified. In the remaining sections I discuss the adverbs in more detail, pointing out some unexpected differences among them.

4.1. Times and time adverbs

The previous example shows that all A-adverbs can measure the event time of a divisible eventuality description, which serves as the basis for classifying these adverbs as A-adverbs. With respect to other time adverbs, however, A-adverbs pattern differently.

It was shown above that English A-adverbs can measure (at least) the event, reference, iterative and habitual time if the appropriate predicate of times is divisible. The reference time can be modified if the reference time predicate contains negation, a downward entailing quantifier, a constituent modified by *only*, or if the viewpoint aspect is imperfective. The *-ig* adverbs can modify the reference time in all of these cases:

(40) (a) János *tíz percig* ment le a lépcsőn (imperfective)
J-nom ten minute-until went down the stair-on
'János was going down the stairs for ten minutes'

(b) János *fél óráig* nem érkezett meg (negation)
J-nom half hour-until not arrived perf
'János didn't arrive for half an hour'

(c) *Fél óráig* kevesebb mint három vendég érkezett meg (decreasing argument)
half hour-until fewer than three guest-nom arrived perf
'For half an hour fewer than three guests arrived'

(d) *Fél óráig* csak János érkezett meg (only)
half hour-until only J-nom arrived perf
'For half an hour only János arrived'

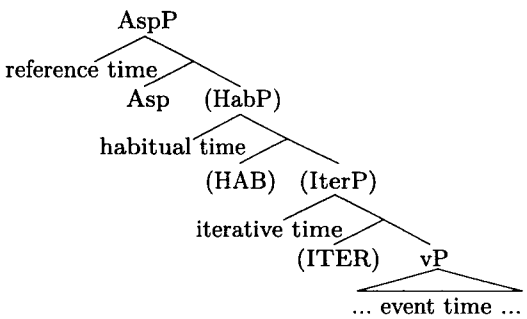
Iterative and habitual times can also be measured by *-ig* adverbs, as shown below. The adverbs measure the time span during which the running or blinking event occurred habitually or iteratively, respectively. Hungarian *-ig* adverbs thus show the same flexibility of time interval modification as English *for*-adverbs.

(41) János *három évig* futott (habitual)
J-nom three year-until ran
'János ran for three years'

(42) A lámpa *tíz percig* pislogott (iterative)
the lamp-nom ten minute-until blinked
'The lamp blinked for ten minutes'

Let us briefly consider the structural position of the time intervals mentioned. I assume that the event time is an argument of the verb, and is merged within the vP. Iterative and habitual interpretations of an eventuality arise as the consequence of the presence of an ITER or HAB operator, respectively. These operators are merged above the vP, and take a time interval argument that is interpreted as the iterative or habitual time.[21] Finally, I assume that the reference time is an argument of the head Asp. Asp contains either a perfective or an imperfective head, encoding the relevant aspectual distinction. The proposed structure, with details omitted, is given below.

(43)



The time intervals in question can all be modified by an *-ig* adverb, as the previous examples show. This flexibility does not extend to all A-adverbs, though; Hungarian A-adverbs differ in the range of time adverbs that they can measure.

[21] I assume that the two operators are distinct and that both may be present in the structure (as in *János coughed for ten years*, for instance). For a discussion of these operators and interpretations, see Carlson (1977); Filip–Carlson (1997); de Swart (1998; 2000); and Rimell (2004), among others

Reference time can only be modified by an *-ig* adverb. Neither of the remaining three adverbs (*át*, *keresztül* or accusative adverbs) can measure the duration of the reference time.[22]

(44) (a) János ??*két órán át* / #*két órán keresztül* / #*két órát*
J-nom two hour-on across two hour-on through two hour-acc
nem érkezett meg
not arrived perf
'János didn't arrive for two hours'

(b) ??*Másfél órán át* / #*másfél órán keresztül* / #*másfél*
one.and.half hour-on across one.and.half hour-on through one.and.half
órát kevesebb mint három vendég érkezett meg
hour-acc fewer than three guest-nom arrived perf
'For an hour and a half, fewer than three guests arrived'

(c) János ??*tíz percen át* / #*tíz percen keresztül* / #*tíz percet*
J-nom ten minute-on across ten minute-on through ten minute-acc
ment le a lépcsőn
went down the stair-on
'János was going down the stairs for ten minutes'

It was shown above that *-ig* adverbs can measure iterative and habitual times as well. The remaining A-adverbs show variable behavior in this respect. Both *át* and *keresztül* can modify these times, while accusative adverbs can modify only iterative, but not habitual time intervals.

(45) (a) János *három éven át* / *három éven keresztül* /
J-nom three year-on across three year-on through
??*három évet* futott
three year-acc ran
'János ran for three years'

(b) A lámpa *tíz percen át* / *tíz percen keresztül* /
the lamp-nom ten minute-on across ten minute-on through
(?)*tíz percet* pislogott
ten minute-acc blinked
'The lamp blinked for ten minutes'

[22] These A-adverbs can give rise to the (irrelevant) reading where the arrival or the application process lasts as long as specified by the adverb. In this case, however, the adverb modifies the event time and not the reference time.

The range of time intervals modified by each A-adverb is summarized below.

(46)

	keresztül	*át*	*-ig*	accusative
event time	OK	OK	OK	OK
iterative time	OK	OK	OK	OK
habitual time	OK	OK	OK	*
reference time	*	??	OK	*

In order to account for the variation observed, I assume that the time interval arguments of adverbs are restricted depending on the position where the adverb is generated or externally merged (as discussed, among others, in Thompson 1996). Time adverbs can only measure the duration of a time interval that is local to the merge position of the adverb. Thus in order for an A-adverb to modify the reference time, it must be merged locally to the reference time; a different, lower local external merge site is required for habitual time modification, and so on. The different behavior of A-adverbs can be encoded by assuming that the external merge position of these adverbs is constrained in different ways. Accusative A-adverbs can be merged locally to vP and IterP. *Keresztül* and *át* can be merged locally to vP, IterP or HabP. Finally, *-ig* adverbs show four-way ambiguity in the position where they are merged; they can be merged locally to vP, IterP, HabP or AspP.

The correlation between the surface position of Hungarian adverbs and their interpretation is consistent with the previous locality generalization. In a negated instantaneous eventuality description like *János didn't arrive*, in (47), a postverbal A-adverb is marked, since it is interpreted as modifying the event time. The reading where the adverb modifies the reference time becomes possible if the adverb is merged higher and precedes the verb.

(47) (a) ??Nem érkezett meg János *másfél* *óráig*
not arrived perf J-nom one.and.half hour-until
'János didn't arrive for an hour and a half'

(b) *Másfél* *óráig* nem érkezett meg János
one.and.half hour-until not arrived perf J-nom
'For an hour and a half, János didn't arrive'

The interpretation of the A-adverb in (48) shows a similar distribution. The postverbal adverb is interpreted as determining the duration of the

sleeping event. A preverbal adverb, in contrast, modifies the reference time and asserts that there was a ninety-minute interval during which János did not sleep.[23]

(48) (a) Nem aludt János *másfél* *óráig*
not slept J-nom one.and.half hour-until
'János didn't sleep for an hour and a half'

(b) *Másfél* *óráig* nem aludt János
one.and.half hour-until not slept J-nom
'János didn't sleep for an hour and a half'

The interaction of A-adverbs and time intervals must be stipulated for each adverb.[24] This restriction can be encoded in various ways; for instance, as part of the definition of the adverb, or as arising from multiple entries for adverbs which can be merged in cartographically distinct positions (possibly as specifiers of different Asp heads) in the structure (Cinque 1999).

(49) The difference in temporal modification among A-adverbs, in terms of the time intervals modified, reduces to the different possible external merge positions of the adverbs.

In addition to the varying interaction with time intervals, A-adverbs show some further differences upon closer inspection. Some of these differences are discussed below.

4.2. Further differences among adverbs

First, *-ig* adverbs are discussed, which show the greatest flexibility in modifying time intervals. A brief description of *keresztül* and accusative adverbs follows.

[23] If the preverbal adverb is a contrastive topic then it can still modify the event time (K. É. Kiss, p.c.). This is consistent with the hypothesis that first merge position determines the constituents modified by the adverb, and topicalization fails to affect that interpretation.

[24] Arosio (2003) notes that Italian A-adverbs have distinct, non-overlapping distributions. That distribution permits a treatment where the distribution is determined by the semantic properties of the complex predicate containing the adverb. This approach is not viable for Hungarian, given the distribution of the A-adverbs as discussed above.

4.2.1. A punctual adverb in disguise

The A-adverb *-ig* differs from other time adverbs in that its measure argument is not necessarily durative.[25] If *-ig* appears with a punctual argument rather than a durative measure phrase, then it establishes the right boundary of a time interval. Since the phrase *két óra* is ambiguous between the two interpretations, *két óráig* is also ambiguous: it is either a punctual or a durative adverb.

(50) János két óráig aludt
J-nom two hour-until slept
'János slept for two hours' (*két óra* = 'two hours' (durative))
'János slept until two o'clock' (*két óra* = 'two o'clock' (punctual))

Even though this characterization highlights the ambiguity of *-ig*, phrasing the distinction this way is deceptive. Durative *-ig* does not, in fact, measure the duration of a time interval. Rather, it can be seen as specifying the endpoint of a time interval, building on the initial point and the duration elapsed since the initial point. This treatment permits a uniform characterization of *-ig* and does not require the assumption of some accidental homonymy.

The punctual *-ig*, where the adverb appears with a punctual argument, does not restrict the eventuality description modified; telic and atelic eventuality descriptions can equally be modified by this adverb.

(51) (a) János 12:30-ig aludt
J-nom 12:30-until slept
'János slept until 12:30'

[25] Despite initial appearances, *until* adverbs show a behavior significantly different from *-ig* adverbs. In affirmative environments, *until* appears with punctual times and modifies only distributive predicates of times. The restriction to distributive event times disappears in the presence of negation—*until*, with a punctual time argument, can modify telic and atelic eventuality descriptions alike (Karttunen 1974; Mittwoch 1977; Giannakidou 2002; among others). First of all note that the variable behavior of *until* arises with punctual time expressions, unlike the variation found with *-ig*. In the case of *-ig*, the different restrictions on the divisibility of the predicates of times depend on the temporal argument of *-ig* (punctual or durative), and are independent of the presence of negation or other downward entailing environments. Furthermore, the different behavior of *until* is not surprising if it is assumed that in negative sentences, *until* modifies the (necessarily divisible) reference time rather than the event time.

(b) János 12:30-ig meg érkezett
J-nom 12:30-until perf arrived
'János arrived by 12:30'

I suggest that punctual *-ig* establishes the right boundary (RB) of a time interval argument of a predicate of times:

(52) 12:30-ig = $\lambda P.\lambda t.[P(t) \,\&\, RB(t)(12:30)]$

The predicate of times and the time interval measured varies according to the properties of the eventuality description. If the eventuality description is atelic, as in (51a) above, then the right boundary established by the adverb is that of the event time. In other words, the sleeping eventuality continued (at least) up to 12:30. Whenever the eventuality description is telic, as in (51b), the time interval modified by *-ig* is the reference time. The eventuality culminated (and therefore the event time ended) before the time established by the adverb.[26] At this point I merely note the difference, and offer no account for it. In addition to event times, punctual *-ig* can also establish the right boundary of iterative, habitual and reference times.

Building on the definition of punctual *-ig*, I suggest that durative *-ig* also establishes the right boundary of a time interval. Two major differences with respect to punctual *-ig* are (a) the restriction of the adverb to divisible predicates of times and (b) the presence of a (possibly implicit) punctual time argument t' which establishes the left boundary (LB) of the time interval. The original definition of an A-adverb, where the measure phrase directly determines the duration of the time interval, is repeated below for convenience.

(53) húsz percig 'for twenty minutes' = $\lambda P.\lambda t.\lambda t'.[\mathbf{LB}(t)(t')$ &
$[\forall t'' \subset t[\exists t'''[t'' \subseteq t''' \subset t \,\&\, P(t''')]]$ & $\mathbf{RB}(t)(t' + [\textbf{twenty minutes}])]$

(54) for twenty minutes = $\lambda P.\lambda t.[\forall t' \subset t[\exists t''[t' \subseteq t'' \subset t \,\&\, P(t'')]] \,\&\, |t| = 20\,\text{minutes}]$

To encode the restriction of durative *-ig* adverbs to divisible predicates of times, I appeal to the definition adopted in the preceding section. The left boundary of the time interval of *-ig* is established by the punctual

[26] The different interpretations of punctual *-ig* are independent of the duration of the eventuality description. A telic event must culminate before the time of punctual *-ig* even if the event description is durative.

argument t', and the right boundary is determined by the left boundary and the duration specified. This treatment of *-ig* adverbs allows a similar treatment of its uses, since *-ig* always determines the right boundary of a time interval. A durative *-ig* adverb thus crucially differs from other A-adverbs, which establish duration rather than the endpoint of the relevant time interval.[27]

4.2.2. Regular spacing of gaps

With respect to the time intervals that they can measure, it was shown that *keresztül* and *át* pattern identically. Both can modify event, habitual and iterative times, but not the reference time interval. I assume that this restriction follows from the stipulated merger sites of the adverbs. Even though their distribution is the same in terms of the time intervals modified, the semantic import of the two adverbs is different. To highlight the difference, let us consider a few scenarios and how the adverbs interact with these.

For a habitual eventuality description, let us assume that János has been learning to play the saxophone for ten years. In the first scenario, he studied and practiced regularly; say, twice a week for ten years. In this case, either an *át* or a *keresztül* adverb can measure the duration of the habitual time. In the second scenario, he takes lessons only intermittently and a number of months can pass without taking a lesson or practicing at all. Here modification of the habitual time by an *át* adverb is more felicitous than by the marked *keresztül.*

(55) János *tíz éven át* tanult szaxofonozni
J-nom ten year-on across learned saxophone-V-inf
'János learnt to play the saxophone for ten years' (regularly/intermittently)

[27] The suffix *-ig* also has a spatial goal interpretation, where it yields the endpoint of a path:

János a házig futott
J-nom the house-to ran
'János ran to the house'

In both spatial and temporal uses, the *-ig* phrase denotes the right boundary of a path or scale. This view permits a homogeneous treatment of both spatial and temporal interpretation of this suffix. In addition, it may reveal a deeper connection between spatial and temporal modification.

(56) János *tíz éven* *keresztül* tanult szaxofonozni
J-nom ten year-on through learned saxophone-V-inf
'János learned to play the saxophone for ten years'
(regularly/??intermittently)

Both *át* and *keresztül* permit gaps between the time intervals for which the predicate of times holds. Intuitively, while *át* does not require the intervals to be spaced at regular intervals, a constant distance between the time arguments of the predicate is required by *keresztül.*

A similar difference is shown by the pair in (57). If János was working on the house only intermittently, when he had money, then modification by *keresztül* is not felicitous. Again, the adverb requires regularly occurring intervals.

(57) (a) János öt éven át építette a házat
J-nom five year-on across built the house-acc
'János was building the house for five years'
(continuously or possibly intermittently)

(b) János öt éven keresztül építette a házat
J-nom five year-on across built the house-acc
'János was building the house for five years' (continuously)

In addition, the adverbs differ in whether they can cooccur with overt frequency predicates. While the frequency can be overtly modified by the equivalent of *often* or *seldom* if the time adverb is *át*, this is not possible with *keresztül.*

(58) (a) János *tíz éven* *át* gyakran/ ritkán látogatta az anyósát
J-nom ten year-on across often seldom visited the mother.in.law-poss-acc
'For ten years, János often / seldom visited his mother-in-law'

(b) ??János *tíz éven* *keresztül* gyakran/ ritkán látogatta az anyósát
J-nom ten year-on through often seldom visited the mother.in.law-poss-acc
'For ten years, János often / seldom visited his mother-in-law'

The restriction on frequency adverbs coocurring with *keresztül* does not extend to all frequency predicates. Some frequency adverbs that can cooccur with these adverbs are grammatical with *keresztül* as well:

(59) János *két éven* *keresztül* hétfőnként látogatta az anyósát
J-nom two year-on across Monday-on visited the mother.in.law-poss-acc
'For two years, János visited his mother-in-law on Mondays'

The differences between *át* and *keresztül* adverbs are thus the following: *keresztül*, unlike *át*, (a) does not allow modification by overt frequency predicates such as *often* and *seldom*, and (b) requires regular gaps in habitual eventuality descriptions.

If these generalizations are on the right track, then the marked behavior of *keresztül* can be treated as stemming from an incorporated frequency predicate *regularly*. If *keresztül* contains a frequency adverb, then properties (a) and (b) follow straightforwardly. The regularity of gaps follows from the interpretation of the adverb. The unavailability of overt frequency modification can be ascribed to a restriction on frequency modification, if only one of a subset of frequency predicates can modify the eventuality description.[28] I propose that the A-adverb definition of Piñón (1999) can be invoked here to define *keresztül* adverbs. The adverb requires a regular distribution of gaps because frequency adverb *regularly* is present, as shown below.

(60) tíz éven keresztül$_1$ 'for ten years' = $\lambda R \lambda P \lambda e[\exists t[[10 \text{ years}](t) \,\&\, \text{Regularly}(e, t, P)]]$

The definition above is true for only those *keresztül* adverbs that apply to the iterative or habitual time. When *keresztül* modifies the event time, no implicit frequency adverb is present. Rather, the adverb shares the definition with other A-adverbs:

[28] More remains to be said about the incompatibility of some frequency predicates and *keresztül*. These overt frequency adverbs can cooccur and show scope interaction:

(i) János *gyakran ritkán* látogatta az anyósát
J-nom often seldom visited the mother.in.law-poss-acc
'János often seldom visited his mother-in-law'
(= it often happened that (during a certain interval) János seldom visited her)

(ii) János *ritkán* látogatta *gyakran* az anyósát
J-nom seldom visited often the mother.in.law-poss-acc
'János seldom visited his mother-in-law often'
(= it seldom happened that (during a certain interval) János visited her often)

If *keresztül* adverbs contain a covert *regularly*, then they are expected to cooccur with frequency predicates such as *seldom, often, from time to time,* etc. It may be also expected that *regularly* takes narrow scope with respect to these predicates. I leave a more detailed discussion of this issue (including deciding the question of whether multiple frequency predicates are possible at all and accounting for the presence or absence of the different readings) for future research.

(61) két órán keresztül$_2$ 'for two hours' = $\lambda P.\lambda t.[\forall t' \subset t[\exists t''[t' \subseteq t'' \subset t \,\&\, P(t'')]] \,\&$
$\& \,|t| = 2\,\text{hours}]$

To summarize: in order to account for the different interpretations of *keresztül*, it was necessary to assume two entries for the adverb. *Keresztül*$_1$ incorporates the frequency predicate *regularly* and modifies iterative or habitual times, where the frequency predicate can be interpreted. The other entry, *keresztül*$_2$, is restricted to modifying the event time. *Át* shows no unexpected behavior. I assume that this is due to the default A-adverb definition of *át*, illustrated in (61).

4.2.3. An accusative adverb

Of the four A-adverbs, accusative adverbs have the most restricted distribution. As noted earlier, they can only modify the event and iterative times, but not habitual or reference time. The modification of habitual and iterative times is illustrated below.

(62) (a) #János *két évet* futott (habitual)
J-nom two year-acc ran
'János ran for two years'

(b) János *fél órát* köhögött (iterative)
J-nom half hour-acc coughed
'János coughed for half an hour'

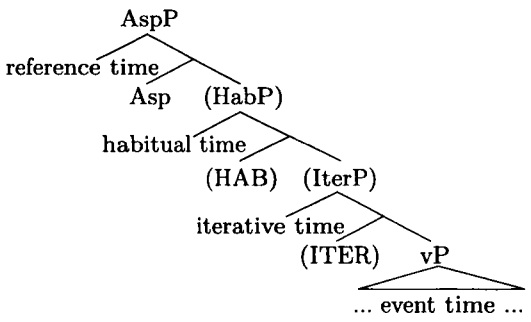
Since the interpretation of accusative and *át* adverbs, apart from the range of time intervals modified, is similar, I assume that accusative adverbs also share the default definition of A-adverbs.

The restriction of accusative adverbs to event and iterative times, similarly to other time interval restrictions, must be independently stipulated. In contrast with this assumption, Csirmaz (2005; to appear) argues that accusative adverbs are excluded from reference time modification because of the need to license the case marking.[29] Accusative case licensing requires a first merge position local to the case licensor v head. Since the adverb is merged low, it cannot modify the reference time, which is merged as the specifier of the higher Asp head.[30]

[29] The proposal is extended to bare time adverbs, which are also assumed to be case marked.

[30] Morzycki (2004) independently suggests a similar restriction on accusative and bare adverbs.

(63)



In order for this account to derive the full range of interpretational properties of accusative adverbs, it needs to be shown that the position of the iterative operator ITER is below that of the habitual operator HAB. Crucially, it must also be shown that ITER is sufficiently close to the v head to permit case licensing of the accusative adverb, while this is not possible if the adverb modifies the habitual time. I leave exploring this possibility for future research, and assume here that the restriction of accusative adverbs is independently stipulated.[31]

4.3. Towards a typology of A-adverbs

The discussion of Hungarian A-adverbs started out with a preliminary definition of A-adverbs, based on Hinrichs (1985) and von Fintel (1997) and repeated below.

(64) for twenty minutes = $\lambda P.\lambda t.[\forall t' \subset t[\exists t''[t' \subseteq t'' \subset t \,\&\, P(t'')]] \,\&\, |t| = 20\,\text{minutes}]$

[31] Maintaining the case-based account seems ultimately unfeasible. Crosslinguistically, accusative and bare adverbs are excluded from modifying the result time, as illustrated below for English bare adverbs.

(i) János opened the window *for two hours*
(ii) #János opened the window *two hours*

A locality account based on case licensing fails to derive this restriction. In addition, the range of time intervals measured by accusative and bare adverbs is not universally constrained to event time and iterative time, as the following Dutch example shows (M. den Dikken, p.c.).

(iii) Fred heeft twintig jaar (lang) niets dan kabeljauw gegeten
Fred has twenty year long nothing but cod eaten
'Fred ate nothing but cod for twenty years'
(habitual time modification)

I suggested that this definition holds for *át* and accusative adverbs. *Keresztül* adverbs are ambiguous. The definition in (64) applies to event time modification. A different definition, repeated below, incorporates the frequency predicate *regularly* and applies to iterative and habitual times.

(65) tíz éven keresztül$_1$ 'for ten years' = $\lambda R \lambda P \lambda e[\exists t[[10 \text{ years}](t) \& \text{Regularly}(e, t, P)]]$

Finally, it was suggested that *-ig* adverbs establish the right boundary of a time interval rather than measure the duration of the interval. In this respect, an *-ig* adverb with a durative complement resembles *-ig* with a punctual time, since both establish the right boundary of some interval complement.

It was also noted that A-adverbs show a variable behavior in terms of the time intervals they can modify. I proposed that the specific range of times be stipulated for each adverb. It is worth pointing out that there is a correlation among the time intervals modified. If an adverb can modify a time interval merged in the structure, then it can modify all of the time intervals merged below that point. *-ig* adverbs, for instance, can modify all four time intervals: reference, habitual, iterative and event time as well. *Keresztül* and *át* adverbs may modify either habitual, iterative or event time. Finally, accusative adverbs can only modify the two lowest time intervals.

5. T-adverbs

The scope of the present section is significantly smaller than that of the preceding discussion, since there are only two Hungarian equivalents of *in*-adverbs: *alatt* and *belül* adverbs.

(66) (a) János *másfél óra alatt* el futott a boltba
J-nom one.and.half hour under away ran the store-to
'János ran to the store in an hour and a half'

(b) János *másfél órán belül* el futott a boltba
J-nom one.and.half hour-on inside away ran the store-to
'János ran to the store in an hour and a half'

One of the striking differences of English *in*-adverbs and their Hungarian counterparts is that Hungarian T-adverbs cannot modify the event time of an imperfective eventuality description.

(67) (a) For two weeks, János was writing the novel in a month
(but then he realized that he'd only finish in three months)

(b) #Két hétig János *egy hónap alatt / egy hónapon belül*
two week-until J-nom one month under one month-on inside
'For two weeks, János was writing the novel in a month'
írta a regényt
wrote the novel-acc
'For two weeks, János was writing the novel in a month'

The constraint on Hungarian T-adverbs can be accounted for by assuming that these adverbs impose a restriction on both the situation and viewpoint aspect properties of the eventuality description that they modify. Both *alatt* and *belül* adverbs modify only perfective telic eventuality descriptions.[32]

Even though both T-adverbs impose identical requirements on the aspectual properties of the eventuality description, they measure distinct time intervals. *Alatt* adverbs measure the duration of the event time and *belül* adverbs, that of the reference time. The difference in the time intervals modified can be shown in a number of environments. Consider, for instance, how T-adverbs affect the interpretation of an instantaneous eventuality description.

If the event time is modified by an *alatt* adverb, the eventuality is interpreted as durative (and has a preparatory stage of some duration), as required by the adverb. In (68a), for instance, János started buying lottery tickets two years before he won the prize. If the reference time is modified by a *belül* adverb, as in (68b), then there is no preparatory stage that would be interpreted as lasting two years. In fact, there is no requirement that a preparatory stage exist at all. The description merely asserts that there was an event of János winning the grand prize at some point within the two-year long interval.

[32] The adverbs restrict the aspectual properties of the eventuality description and not divisibility. This is shown by grammaticality of *alatt* adverb modification of the event time below:

János (két napig) nem válaszolt tíz perc alatt
J-nom two day-until not answered ten minute under
'For two days, Janos didn't answer in ten minutes'

In this case the reference time predicate is divisible (as in the case of imperfective eventuality descriptions), but the viewpoint aspect is perfective— as required by the adverb.

(68) (a) János *két év alatt* meg nyerte a főnyeremény
J-nom two year under perf won the grand.prize-acc
'János won the grand prize in two years'

(b) János *két éven belül* meg nyerte a főnyeremény
J-nom two year-on inside perf won the grand.prize-acc
'János won the grand prize in two years'

The same difference is shown by the following example. The *alatt* adverb, which modifies the event time, enforces a marked interpretation with a preparatory stage. The resulting interpretation is that Juli was actively engaged in trying to stumble for ten minutes, which she managed to do at the end of the interval. With a *belül* adverb modification, in contrast, the eventuality description merely asserts that Juli stumbled within a ten-minute interval.

(69) (a) ??Juli *tíz perc alatt* meg botlott
J-nom ten minute under perf stumbled
'Juli stumbled in ten minutes'

(b) Juli *tíz percen belül* meg botlott
J-nom ten minute-on inside perf stumbled
'Juli stumbled in ten minutes'

The difference between the two T-adverbs can also be shown with durative predicates. As noted earlier, the reference time of perfective eventuality descriptions properly contains the event time. Thus if the duration of an event is explicitly specified, only *alatt* adverbs can measure that duration. As an illustration, let us consider a situation where János starts writing a letter at 4 o'clock and finishes exactly at 5 o'clock.

(70) (a) János (*pontosan*) *egy óra alatt* meg írta a levelet
J-nom exactly one hour under perf wrote the letter-acc
'János wrote the letter in an hour'

(b) #János (*pontosan*) *egy órán belül* meg írta a levelet
J-nom exactly one hour-on inside perf wrote the letter
'János wrote the letter in an hour'

In the situation where the event time lasts exactly 60 minutes, modification by an *alatt* adverb is grammatical, but a *belül* adverb is not felicitous. This follows if the reference time must properly include the

event time; the reference time interval must measure longer than sixty minutes in this case.

To conclude this section, let us consider the definitions and restrictions of Hungarian T-adverbs. Both *alatt* and *belül* adverbs, just as English *in*-adverbs, can be defined as given in (71).

(71) in twenty minutes = $\lambda P.\lambda t.[\neg[\forall t' \subset t[\exists t''[t' \subseteq t'' \subset t \,\&\, P(t'')]]] \,\&\, |t| = 20\,\text{minutes}]$

In addition, it must be specified that the two adverbs measure different time intervals: *alatt* adverbs measure the duration of the event time and *belül* adverbs, that of reference time. Furthermore, it must be ensured that both adverbs modify only perfective telic eventuality descriptions. While the non-local restriction of T-adverbs on predicates of times is puzzling—with *alatt* requiring perfective aspect (a property of Asp) and *belül* requiring telic aspect (a property of vP)—it is a restriction that must nevertheless be incorporated into an account of these adverbs.

6. Time adverbs in Hungarian

To conclude, let us review the main claims of the paper. First, it was noted that time adverbs can modify not only the event time, but also the iterative, habitual time or the perfect time span, or the reference time. Building on Hinrichs (1985) and von Fintel (1997), I suggested preliminary definitions for Hungarian time adverbs—which serve as the definitions of English time adverbs— that accommodate the possibility of modifying these distinct times.

A discussion of Hungarian time adverbs revealed that the adverbs are not freely interchangeable. Time adverbs (both A- and T-adverbs) differ in the time intervals they can modify. I suggested that this difference should be explicitly encoded as an arbitrary restriction on the adverbs. I showed that Hungarian T-adverbs only differ in the time intervals they measure, but otherwise impose identical restrictions on the eventuality description they modify. I also suggested that for Hungarian A-adverbs, the time intervals that they can measure must be specified and does not follow from independent factors.

Two divergences from the default A-adverb definition were also identified. First, in order to ensure a uniform treatment of *-ig* adverbs, I proposed that durative *-ig* adverbs establish the right boundary of the time interval they modify, similarly to punctual *-ig* adverbs. With durative

-ig adverbs then, the measure phrase measures the distance between the left and right boundaries of the time interval. For *keresztül* adverbs, I suggested that they are ambiguous. The default definition applies if the adverb measures the event time. For iterative and habitual time modification, in contrast, a different denotation was invoked, where the adverb contains an implicit frequency predicate *regularly*.

This preliminary study of Hungarian time adverbs reveals a number of consequences for the theory of adverbial modification. It appears necessary to impose explicit restrictions on the range of time intervals that can be modified by specific time adverbs; the distinct possibilities of adverbial modification do not follow readily from independent considerations. In addition, adverbs may impose non-local restrictions on semantic properties: Hungarian T-adverbs constrain both situation and viewpoint aspectual properties, independently of the time intervals they measure. Finally, the discussion suggests that there is no unique definition of time adverbs. While most adverbs conform to a default view of adverbs — one that is based on the notion of divisibility of Hinrichs (1985) and von Fintel (1997) — not all adverbs do so. I suggested that the exceptions are Hungarian *-ig* and *keresztül* adverbs. The uniform treatment of the spatial and temporal uses of *-ig* adverbs (which warranted an unorthodox view of measuring duration) suggests that the parallels between the two domains of modifications may extend yet further. An exploration of this connection and the issues noted above is left for further research.

References

Arosio, Fabrizio 2003. Temporal homogeneity and the Italian perfect. In: Artemis Alexiadou – Monika Rathert – Armin von Stechow (eds): Perfect explorations, 37–66. Mouton de Gruyter, Berlin & New York.

Bach, Emmnon 1986. The algebra of events. In: Linguistics and Philosophy 9 : 5–16.

Bennett, Michael – Barbara Partee 1972. Toward the logic of tense and aspect in English. System Development Corporation, Santa Monica CA. (Reprinted in Partee, Barbara: Compositionality in formal semantics, Malden MA & Oxford, Blackwell, 2004.).

Bertinetto, Pier Marco 2001. On a frequent misunderstanding in the temporal–aspectual domain: The "perfective–telic" confusion. In: Carlo Cecchetto – Gennaro Chierchia – Maria Teresa Guasti (eds): Semantic interfaces. Reference, anaphora and aspect, 177–210. CSLI Publications, Standford CA.

Carlson, Gregory 1977. Reference to kinds in English. Doctoral dissertation, University of Massachusetts, Amherst.

Cinque, Guglielmo 1999. Adverbs and functional heads: A cross-lingustic perspective. Oxford University Press, New York & Oxford.

Comrie, Bernard 1976. Aspect. Cambridge University Press, Cambridge.

Csirmaz, Anikó 2005. Semantics and phonology in syntax. Doctoral dissertation, MIT.

Csirmaz, Anikó 2006. Aspect, negation and quantifiers. In: Katalin É. Kiss (ed.): Event structure and the left periphery, 225–54. Springer, Dordrecht.

Csirmaz, Anikó to appear. Measuring time. In: Proceedings of the Penn Linguistics Colloquium 30. University of Pennsylvania Working Papers in Linguistics, Philadelphia.

Demirdache, Hamida – Myriam Uribe-Etxebarria 2000. The primitives of temporal relations. In: Roger Martin – David Michaels – Juan Uriagereka (eds): Step by step. Essays in honor of Howard Lasnik, 157–86. MIT Press, Cambridge MA.

Demirdache, Hamida – Myriam Uribe-Etxebarria 2004. The syntax of time adverbs. In: Jacqueline Guéron – Jacqueline Lecarme (eds): The syntax of time, 143–80. MIT Press, Cambridge MA.

Dowty, David 1979. Word meaning in montague grammar. Dordrecht, Reidel.

É. Kiss, Katalin in press. Jól megoldottuk? Rosszul oldottuk meg? [Have we solved it correctly? Have we solved it incorrectly?]. In: Magyar Nyelv 102.

Filip, Hana – Gregory Carlson 1997. Sui generis genericity. In: Penn Working Papers in Linguistics 4:91–110.

Fintel, Kai von 1997. Aspectual classes and aspectual composition. Handout, MIT.

Fintel, Kai von 1999. NPI licensing, Strawson entailment, and context-dependency. In: Journal of Semantics 16:97–148.

Fintel, Kai von – Sabine Iatridou 1997. Since since. Manuscript, MIT.

Giannakidou, Anastasia 2002. UNTIL, aspect and negation: A novel argument for two untils. In: Brandon Jackson (ed.): Semantics and linguistic theory (SALT) 12, 84–103. CLC Publications, Cornell University, Ithaca.

Hinrichs, Erhard 1985. A compositional semantics for Aktionsarten and NP reference in English. Doctoral dissertation, Ohio State University.

Iatridou, Sabine – Elena Anagnostopoulou – Roumyana Izvorski 2001. Some observations about the form and meaning of the perfect. In: Michael Kenstowicz (ed.): Ken Hale: A life in language, 189–238. MIT Press, Cambridge MA. (Reprinted in Artemis Alexiadou – Monika Rathert – Arnim Stechow (eds): Perfect Explorations, 153–204. Mouton de Gruyter, Berlin, 2003.).

Karttunen, Lauri 1974. Until. In: Chicago Linguistic Society 10:284–97.

Klein, Wolfgang 1994. Time in language. Routledge, London & New York.

Krifka, Manfred 1998. The origins of telicity. In: Susan Rothstein (ed.): Events and grammar, 187–235. Kluwer, Dordrecht.

Mittwoch, Anita 1977. Negative sentences with *until*. In: Chicago Linguistic Society 13:410–7.

Moltmann, Frederike 1991. Measure adverbials. In: Linguistics and Philosophy 14: 629–60.

Morzycki, Martin 2004. Measure DP adverbials: Measure phrase modification in VP. Mansucript, UQAM.

Olsen, Mari Broman 1997. A semantic and pragmatic model of lexical and grammatical aspect. Garland, New York.

Piñón, Christopher 1999. A semantics for durative adverbials. Paper presented at Sinn und Bedeutung, Düsseldorf, 1999.

Rimell, Laura 2004. Habitual sentences and generic quantification. In: Vineeta Chand – Ann Kelleher – Angelo Rodriguez – Benjamin Schmeiser (eds): Proceedings of WCCFL 23, 663–76. Cascadilla Press, Somerville MA.

Rothstein, Susan 2004. Structuring events. Blackwell, Malden MA & Oxford.

Smith, Carlota 1991. The parameter of aspect. Kluwer, Dordrecht.

Stechow, Arnim von 2002. Temporal prepositional phrases with quantifiers. Some additions to Pratt and Francez. Manuscript, Universität Tübingen.

Swart, Henriette de 1998. Aspect shift and coercion. In: Natural Language and Linguistic Theory 16:347–85.

Swart, Henriette de 2000. Tense, aspect and coercion in a cross-linguistic perspective. In: Miriam Butt–Tracy Holloway King (eds): Proceedings of the Berkeley Formal Grammar Conference. University of California, Berkeley. (Online publication at http://csli-publications.stanford.edu/LFG/5/bfg00/bfg00deswart.pdf)

Thompson, Ellen 1996. The syntax of tense. Doctoral dissertation, University of Maryland at College Park.

Vendler, Z. 1967. Linguistics in philosophy. Cornell University Press, Ithaca.

Zucchi, Sandro 1991. Negation and aspect. In: Youngkyoon No – Mark Libucha (eds): Proceedings of the Eighth Eastern States Conference on Linguistics, 344–63. CLC Publications, Ithaca.

Acta Linguistica Hungarica, Vol. 53 (3), pp. 291–332 (2006)
DOI: 10.1556/ALing.53.2006.3.3

PREDICATE FRONTING AND DATIVE CASE IN HUNGARIAN*

BARBARA ÜRÖGDI

Research Institute for Linguistics
Hungarian Academy of Sciences
Benczúr utca 33.
H–1068 Budapest
Hungary
urogdi@nytud.hu

Abstract: In this paper I provide a unified analysis of predicate clefts and a large class of secondary predicates in Hungarian. These constructions involve the occurrence of a predicate that is not immediately dominated by tense, which results in a striking similarity: in both, a verbal predicate takes the form of an infinitive, while nominal/adjectival predicates appear as dative. I argue that both the infinitive and the dative surface form indicate that the predicate's head is spelled out in a functional projection (commonly referred to as PredP or AspP) just outside the predicate's lexical projection, which is evidenced, among other facts, by phrase-internal modification patterns. Therefore, the structure of predication, including the position of modifiers, is claimed to be uniform, regardless of the lexical category of the predicate itself (V, A or N). This result is strongly in line with research that views datives as predicates in general, and structural case on nominals as a reflex of a functional projection in the tense-aspect domain.

Keywords: dative, infinitive, predicate cleft, predication, secondary predication

Introduction

In this paper I provide a unified analysis of the predicate cleft construction and secondary predication in Hungarian. I argue that these two constructions have a crucial property in common — namely, the occur-

* I am grateful first and foremost to Katalin É. Kiss and Marcel den Dikken for invaluable discussions on this topic.

rence of a predicate that is not dominated by its own tense projection—that explains some interesting similarities between them. Although this structural analogy is evident in many languages when it comes to verbal predicates (which are usually infinitival in predicate clefts as well as in secondary predication), I will concentrate here also on nominal and adjectival predicates because in Hungarian these demonstrate a similar parallelism: the predicate occurs with dative case in both predicate clefts and in secondary predication. Based on the shared properties of secondary predication and predicate fronting on one hand, and the common distribution of infinitives and datives on the other hand, I argue for the following:

1. The structure of the predicative phrase, including the position of modifiers, is uniform, regardless of the lexical category of the predicate X itself (V, A or N).
2. At least in Hungarian, datives and infinitives are both the realization of an X head spelled out in the head of a functional projection dominating the lexical projection XP.
3. Therefore, the programmatic claim that datives are predicates receives strong support from the facts and analysis discussed here.

Unlike verbal predicate clefts, adjectival/nominal predicate fronting has not received much attention in the literature (although see Burányi 2003).[1] An illustrative example is given below:

(1) **Gazdag-nak´** (János) **gazdag** volt (de mégsem volt boldog).
rich-dat (John) rich was (but yet-not was happy)
'Rich he (John) was, but he still wasn't happy.'

There are two characteristics of this construction that warrant a closer look. Firstly, it exhibits a well-known property of predicate clefts, namely a mismatch between the fronted predicate and the base predicate—the first being dative, while the latter bears no case. This mismatch manifests itself in verbal predicate cleft constructions as in (2):

(2) **Beszél-ni´ beszél-tem** (de nem figyelt senki).
speak-inf speak-past.1sg (but not listened nobody)
'Speak I did, but nobody was listening.'

[1] In what follows, I use ´ to indicate the typical rising intonation of contrastive topic, to differentiate it from focus, which I indicate with small capital letters.

This evident mismatch is an issue that has mostly been glossed over by the authors on verbal predicate clefts for an obvious reason: intuitively, the topicalized form of the verb, that is, the infinitive, is somehow a "neutral", "unmarked" or "default" form. (Although note that this unmarkedness is only semantic and possibly syntactic—morphologically the infinitive is often not the bare form, as the Hungarian example above shows.) However, it would be difficult to claim that the dative form of an adjective or noun is less marked than the nominative or caseless form. In fact, at first glance, the exact opposite seems to hold.

A second, related issue is the mere fact that a predicate occurs with case. This is not a normal state of affairs, although we know of constructions—from Hungarian as well as other languages—where secondary predicates are case-marked. Incidentally, a large class of secondary predicates in Hungarian also surface as a dative[2] if non-verbal, and as an infinitive if verbal:

(3) (a) János **boldog-nak** látszik.
John happy-dat seems
'John seems happy.'

(b) János **örül-ni** látszik.
John be-happy-inf seems
'John seems to be happy.'

Therefore the main line of my analysis will be the following: It is not an accident that secondary predicates and fronted predicates look the same,[3]

[2] Not all secondary predicates take dative in Hungarian, for example:

(i) János piros-**ra** festette a kerítést.
John red-onto painted the fence
'John painted the fence red.'

(ii) A kerítés piros-**an** tetszik nekem.
the fence red-on appeals to-me
'I like the fence red.'

And so on. These (and a handful of other) suffixes can also mark secondary predicates, and also alternate with a caseless form (*A kerítés piros* 'The fence is red'). Many but not all of these suffixes are related to locative case-endings, so it is plausible that they—like the dative—are spell-outs of a functional head selected by the matrix verb. This is something to be looked at in further research.

[3] This also ties in with an observation in Landau (2004), who credits various authors for it: namely, that whatever constituent takes part in predicate fronting should be independently derivable in the language. While the relevant phrase observes

since they share the property of not being "primary", that is, not being dominated by Tense. (For the basics of the idea that category and other formal properties are determined in syntax based on configurations of lexical and functional heads, see Pesetsky (1995) and others on Distributed Morphology.) I will claim that all of the examples (1)–(3) involve movement of the same phrase, and that the only difference between the doubling and non-doubling (i.e., "secondary predication") cases lies in the rest of the structure: whether or not the same predicate also happens to be the main predicate in the clause.

The paper is organized as follows. In section **1**, I discuss the relevant examples from Hungarian in detail, and show that the construction in (1)–(2) is basically a standard instance of predicate fronting, similar to but less restricted than that found in Russian, Hebrew, Yiddish and a host of other languages. In section **2**, I demonstrate the status of dative case in Hungarian and show that its distribution is in fact very similar to that of the infinitive. I will argue that this parallelism between the dative and the infinitive is well founded syntactically and makes sense semantically as well. Section **3** deals with a question that is relevant to all predicate cleft constructions: whether it can be shown that these constructions involve movement. In section **4**, I demonstrate that predicate (incl. secondary predicate) fronting involves phrasal movement, in particular, the movement of the functional projection dominating AP/NP/VP. Section **5** presents a brief overview of possible alternative analyses of this construction and counterarguments to these. Section **6** contains the summary, as well as an outline of the possible extensions of my conclusions.

1. Data—a detour into Hungarian syntax

In this section I will discuss the relevant observations about predicate fronting in Hungarian. It will be shown that, unlike in many other languages discussed in the literature, there is no uniform position targeted by the movement, no restriction on the surface position of the base copy, and no semantic requirements about the nature of the predicate. Although there exist certain syntactic constraints on the internal structure of the fronted phrase, these are related to general rules of Hungarian syn-

different restrictions in Hungarian from the ones noted for English and Hebrew (etc.), the intuition that **these phrases should not be attributed properties not attested elsewhere in the grammar** holds here as well.

tax and are not unique to this construction. Therefore, the construction can be treated as a simple instance of predicate fronting, without any "special" properties.

As discussed in semantic literature on predicate clefts (among others: McCoy 2002), these constructions usually involve a special interpretation, which is commonly associated with contrastive topics. Observe this in the case of a simple example:

(4) [**Tanár-nak**$^{\nearrow}$] **tanár** vagyok (de nem valami jó).
teacher-dat teacher I-am (but not very good)
'It is true that I'm a teacher but I'm not very good at it.'
Possible scenario: First speaker assumes that second speaker is good with children since he is a teacher. Second speaker concedes the condition (he is in fact a teacher) but contests the entailment (that he should be good with children) on the grounds of a piece of information not known to the first speaker (namely, that he is not good at his job).

It has been claimed or tacitly assumed for some languages that the semantic contrast in the above example is entailed by the predicate cleft construction: The speaker concedes a certain point, but at the same time states that that point is either not relevant, or does not lead to the consequence assumed by the listener/previous speaker. This is true for the typical case in Hungarian (4) as well but not a necessary condition on predicate fronting:

(5) Csak (elég) GAZDAG-NAK nem elég **gazdag** (amúgy tökéletes).
only (enough) rich-dat not enough rich (otherwise perfect)
'It's only that he is not rich enough, otherwise he is perfect.'

Example (5) involves focusing of the predicate (marked by *csak* 'only')—the fronted predicate (in small capitals) occupies focus position, is not a contrastive topic, and accordingly receives focus (and not contrastive topic) interpretation (to be clarified below).

Thanks to the articulated information structure of Hungarian (see e.g., É. Kiss 2002 and her earlier work), topic and focus are distinguished by word order as well as intonation:

(6) (a) **Az elnökkel´** meg-ismerkedtem (de amúgy unalmas volt a parti).
the president-instr perf-I-met (but otherwise boring was the party)
'The President I met, but otherwise the party was boring.'
Assumption: The speaker's goal at the party was to meet the president; therefore, he must have enjoyed himself if he succeeded in his goal.
Assertion: Speaker concedes that he met the president, but contests the entailment—claims that meeting the president did not make the party pleasant.

(b) **AZ ELNÖKKEL** ismerkedtem meg (nem pedig a helyettesével).
the president-instr I-met perf (not conj the deputy-instr)
'It is the president I met, not the vice president.'
Assumption: The speaker met someone (and it was the vice president).
Assertion: It was the president and not anyone else that the speaker met.

As shown by the word order (observe the position of the perfective particle *meg* in the two examples above) and the intonation (rising on *az elnökkel* in (6a) but rising-falling in (6b)), the phrase *with the president* is in (contrastive) topic position[4] in (6a) but in focus position in (6b). This is reflected in the interpretation. Similarly, the interpretations of (4)–(5) derive from the different target positions of the fronted predicate phrase, and not from the fact of the fronting itself.

Just as the position targeted by predicate fronting is not uniform, the surface position of the base copy of the predicate is also not restricted—although it is often focused, this is not obligatory, shown by example (7) below, where another element (the subject) is in focus:

(7) [Gazdag-nak´] PÉTER gazdag (de nekem mégis János tetszik jobban).
rich-dat Peter rich (but I still like John better)
'It's Peter who is rich, but I still like John better.'
Assumption: Speaker likes rich men and would therefore like Peter, not John.
Assertion: Speaker concedes that Peter is the rich one but states that the entailment in the assumption does not hold true.

So the generalization that seems to hold about Hungarian predicate clefts is that the surface positions of the fronted predicate or of the predicate

[4] Whether the **contrastive** topic position is a syntactically unique position or one of the iterative topic projections is a question that I leave open here. While it is true that contrastive topics receive special intonation and semantics, and there can only be one such topic per sentence, they can be preceded and followed by regular topics without restriction. This is not relevant to anything I have to say here.

left behind inside the clause are irrelevant for the availability of fronting and the assignment of dative case.[5] Whenever the predicate is fronted, the doubling and the dative case become available. Therefore—given the fact that there are also no semantic restrictions as to which predicates can be doubled—in what follows I will treat these examples as general cases of predicate fronting, with the added complication of double pronunciation, regarding which I refer the reader to the extensive literature on the topic.[6]

Now let us examine the properties of the fronted predicate phrase. The main constraints on nominal and adjectival predicate fronting are summarized below:

(8) Post-head modifiers are not present in the fronted phrase:[7]
[Büszké-nek (*Péterre)] büszke volt (Péterre),
proud-dat Peter-loc proud was (Peter-loc)
de sajnálta, hogy kevesen látták a győzelmét.
but he-was-sorry that few saw the victory-acc
'As for being proud of Peter, he was, but he was also sorry that so few people had seen his victory.'

(9) The copula is not present in the fronted phrase:
[Büszké-nek (*volt/*lenni)] büszke volt.
proud-dat was/ to be proud was
'As for being proud, he was.'

Notice that this latter fact points to an analysis of the copula that I will assume here without defending it: that the Hungarian copula is not a verb, but rather the spell-out of some functional projection outside the

[5] To position this phenomenon among predicate fronting constructions cross-linguistically: As evidenced by the languages examined in the literature (see references at the end of this paper), predicate fronting constructions fall into two classes. In the first class (seen in various Creole languages, among others) the fronted predicate receives a particular interpretation (i.e., contrastive topic) and there are strong semantic and syntactic restrictions applying to the fronting. In other languages (such as Hebrew or Hungarian) the fronting is much freer, and the syntactic process much more easily viewed as a regular instance of constituent fronting.

[6] For a comprehensive overview of multiple spell-out options in predicate fronting cross-linguistically, as well as strong arguments for the compatibility of multiple pronunciation and movement, see Landau (2006) and references therein.

[7] Whether post-head modifiers are not there in the fronted phrase or they are deleted is a question I return to later on.

predicate phrase that is fronted here (probably T).[8] This is not a general restriction on BE since it can occur in a fronted predicate when it is inserted under a verbal node—shown by the fact that it is modified by an adverb:

(10) [Otthon len-ni] otthon volt (de nem nyitott ajtót).
home be-inf home was (but not opened door-acc)
'As for being at home, she was, but she wouldn't open the door.'

Without further discussion I will just treat examples like (10) as simple verbal examples, where the pre-verbal modifier (often labeled "verb-modifier" in grammars of Hungarian) fronts along with the verb. This is a general constraint on predicate fronting examples involving a verbal predicate, as shown below:

(11) (a) [Haza men-ni] haza ment (de nem tudott aludni).
home go-inf home went (but not could sleep)
'He did go home but he couldn't fall asleep.'

(b) [Jól ír-ni] jól ír (de nincs benne önfegyelem).
well write-inf well he-writes (but there-isn't in-him discipline)
'He does write well, but he doesn't have any discipline.'

(c) [Kez-et ad-ni] kez-et adott (de rögtön továbbállt).
hand-acc give-inf hand-acc he-gave (but immediately went-on)
'He did shake hands with me, but then he immediately went away.'

(d) ?[Péter-t választa-ni] Péter-t választottam
Peter-acc choose-inf Peter-acc I-chose
(de már nem tudom, miért).
(but anymore not I-know why).
'As for choosing Peter, I did, but I no longer know why.'

As the above examples show, basically any complement can front with the verb, as long as it is **left-adjacent to the verb in the surface form of the base sentence**—in the terms of standard Hungarian syntax it is "in the verb-modifier position". Nominal/adjectival predicates are subject

[8] Similar restrictions on the fronting of the copula have been observed for Hebrew (Landau 2004) and Russian (Abels 2001). While my analysis (outlined in what follows) provides a straightforward explanation for the restriction as it applies to Hungarian, it is unclear at this point how and whether it can be extended to cover the Hebrew and Russian facts. I will offer some tentative comments on this matter later on.

to the same restriction when fronted: modifiers of the adjective can front along with the predicate as long as their neutral position is pre-head, as illustrated below—imagine the following sentences in the context of a casting agent describing potential candidates for certain parts:

(12) (a) Átlagos testalkatúnak átlagos testalkatú volt...
average built-dat average built was
'As for having an average build, he did (but the costume still didn't fit him).'

(b) Közepesen magasnak közepesen magas volt...
medium tall-dat medium tall was
'As for being medium height, he was (but he wasn't what we were looking for).'

(c) *Péter-nél magasabb-nak Péter-nél magasabb volt...
Peter-adess taller-dat Peter-adess taller was
'As for being taller than Peter, he was (but otherwise he didn't fit the description).'

I will return to this point in section 4. For the time being, it should be noted that it is an independent fact of Hungarian syntax that a designated, usually predicative element (be it a particle, an adverb, a secondary predicate or an argument) occupies the preverbal position in a neutral sentence.[9] This element is commonly referred to in the literature as the "verb-modifier" (or VM) because it displays a close connection with the verb by modifying its meaning, is often idiomatic (as in (11c) above), and has a fixed syntactic position. A similar restriction[10] governs the neutral position of a modifier to an adjectival or nominal predicate—some modifiers (which, loosely speaking, describe a subtype in the case of a type predicate) are only natural in a pre-head position. While this is not by far a straightforward issue syntactically or semantically, it is marginal to what I have to say here. The only relevant point here is that the

[9] The status of this element is a much debated area of Hungarian syntax. For one, its presence has aspectual relevance. Also, it is difficult if not impossible to distinguish it from focus, leading many researchers—among them myself—to believe that the "VM" position is indistinct from the position occupied by focus. Since this debate is largely internal to Hungarian grammar and would only confuse the non-Hungarian reader, I will attempt to stay as neutral on this matter as my subject allows.

[10] For some reason, the relationship between an adjectival predicate and its modifier has not received as much attention in the literature as the "verb-modifier" has. Nevertheless, I will treat the two types of modification analogously, which will be supported by the facts of predicate fronting.

similar modification of different categories of predicates supports the idea that there is no crucial difference between them at the level of predicative structure.[11] Therefore, I will not dwell too much on **why** this element occurs before the head, and will simply note that **whatever element occurs immediately left of and in the same intonational phrase as the predicative head in the base sentence** can and must front along with its predicate. In what follows, I will refer to such modifiers as **predicate-modifiers** or **PM** for short. For now, I will use this term as descriptive shorthand, although I will offer some thoughts on the nature and syntactic status of PM later in this paper. To allow for simplicity of presentation, I will simply label the functional projection outside the lexical predicative phrase "FP" (as opposed to the more meaningful "PredP" or "AspP" that are used in the literature)—by doing so, I wish to avoid taking a stand on the exact nature of this projection, simply noting that it houses the main assertion in the Hungarian clause. Thus, I will assume that the prehead position of the PM arises via head movement of the lexical predicate to the head of FP and raising the PM from a lower predicative position into [SpecFP], as in (13):

(13) $[_{FP}$ haza$_i$ $[_{F'}$ ment$_j$ $[_{VP}$ t$_j$ ... t$_i$]]]

I take FP thus to be the locus basically of complex predicate formation wherein lower predicates become the modifiers of higher predicates, while at the same time retaining their dominant position by landing on the left edge, and thus receiving main sentence stress. Whether the "low predicative position" is that of a low argument in a Larsonian model, or the predicate of a small clause is not important for the purposes of the present topic. Although I am more sympathetic to the latter approach, discussing it would take me too far off course.

Note that the positional requirement of the PM is so strong that it will be pre-head in the fronted phrase even if this does not mirror the state of affairs inside the base sentence:

(14) [El-lop-ni] PÉTER lopta el a könyvet,
away-steal-inf Peter stole away the book-acc
(de a húga hozta vissza).
(but the sister-his brought back)
'As for stealing the book, that was Peter's doing, but it was his sister who returned it.'

[11] Contra, among others, famous work by Bowers on predication, who claims that only verbal heads move out of their lexical projection to Pr(ed). For the latest version of Bowers' predication theory, see Bowers (2001).

In (14), the particle *el* is in the preverbal position in the fronted phrase, but inside the base sentence it has separated (for independent reasons) from the verb, since the verb must be right-adjacent to the focused element—i.e., there is a functional projection housing the focused subject in its Spec and the verb in its head.[12] Although the internal structure of the base sentence presents a theoretical puzzle in itself, it is at least clear that the surface order in the fronted phrase does not mirror the surface order in the base —ruling out an approach involving PF-echo effects. This is shown also by cases of long-distance fronting, as in (15):

(15) El-lop-ni el akartam [t] (de nem sikerült).
away-steal-inf away I-wanted (but not worked)
'As for stealing it, I wanted to do that but it didn't work out.'

Although this is a complicated example and I will provide a detailed analysis of it in section 4, it is easy to note again that the PM occurs before "its own" predicate in the fronted phrase (*el-lopni*) although such a combination is not found in the rest of the sentence (because, again for independent reasons, *el* has raised up to become the PM of the matrix predicate *akartam* in a form of clause union; on the details of this see É. Kiss 2002).

So, based on the above, the generalization is that the predicate plus its PM are what can (and must) be included in the fronted phrase. As a related fact, subjects can never front:

(16) (a) *[Én utál-ni] én utálom Pétert (de miért fontos ez?)
I hate-inf I hate Peter-acc (but why important this)
Intended meaning: 'It's true that I'm the one who hates Peter...'

(b) ??[Vendég érkez-ni] vendég érkezett (de már el is ment).
guest arrive-inf guest arrived (but already away part went)
Intended meaning: 'It's true that a guest arrived but he's already left.'[13]

[12] Whether this functional projection is a special "FocusP" or not is again an open question. In what follows, I will label it as such for ease of exposition, noting that there are a good number of other theoretical possibilities (i.e., that "FP" housing the PM is iterative, or that sentences containing a Focus are biclausal, with Focus predicated of the entire sentence, and so on).

[13] The latter example is strange because the surface subject is actually a deep theme, which does end up in the preverbal position in a neutral sentence due to a restriction on non-specific bare nominal arguments (i.e., **Érkezett vendég* is ungrammatical)—therefore its inability to occur in the fronted position is unex-

A set of interesting examples involves an added layer of predication on an argument of the verb, usually the direct object, although other arguments are also possible:

(17) (a) **Vers-nek** vers-et írt (de szabadverset).
poem-dat poem-acc she-wrote (but free-verse-acc)
'It's true that it was a poem that she wrote but it was free verse.'

(b) **Vers-nek** vers volt, amit írt...
poem-dat poem was what wrote...

The interpretation of such sentences is similar to that of (4), so they would both be natural in the following situation: The kids in a class are supposed to write poems for Christmas. The teacher of the class is complaining to a colleague about a certain student who never does what she is supposed to. The colleague asks: "So, again she did something other than the task? Did she not write a poem but something else—a joke, a cartoon caption, a short story?" The teacher responds: "It was a poem that she wrote, that's not the problem—the problem is that she wrote free-verse." While (17b) obviously fits in with the other facts I deal with here since it involves a predicate (*poem*) fronted and receiving dative, (17a) is seemingly an exceptional case. It should be noted, however, that all of my examples involve "predicate fronting" of elements that start from a predicative position in the derivation—and it has been argued that all PM's share this property. In particular, Komlósy (1992) claims that a complex predicate such as *verset ír* 'poem-writes' starts from a deep structure like "he writes [something, and that something is a] poem", which would technically make the PM in (17a) a null-headed relative that is predicated of an empty direct object. (See den Dikken forthcoming for a parallel analysis of English predicate inversion.) Although the scope of this paper does not allow for the detailed elaboration of this idea, I

pected. But (16b) is not exactly as bad as (16a), which might mean that (16b) is not ungrammatical but pragmatically or semantically odd. Since subjects in the preverbal position have a tendency to be interpreted as focus in Hungarian (given that the preverbal focus is basically indistinguishable from—and possibly syntactically in the same position as—the PM; see footnote 9) the oddness might be due to trying to interpret the same element "guest" as focus and topic simultaneously. This is, of course, impossible (see for example den Dikken forthcoming for discussion). (See also the slightly marginal status of (11d).) The point is that in (16b) the interpretation must be construed so that *vendég érkezett* 'guest arrived' is a complex predicate, otherwise "guest" must be interpreted as focus, and the sentence becomes incoherent.

believe that it can be convincingly shown that examples (17a) and (17b) are more analogous than the surface shows, and thus my analysis can be extended to both.[14]

Interim conclusion 1:

The properties displayed by the fronted phrase are unrelated to the surface position occupied by the phrase (topic or focus). Therefore I hope to show in what follows that all properties (content and morphology) of the fronted phrase can be derived from its internal structure and its starting position (i.e., that it must be predicative). Whenever and wherever such a fronted phrase is spelled out, it will have the same properties (dative case on adjectival/nominal predicates, and verbal predicates appearing in the infinitival form, each accompanied by its normal predicate modifier).

[14] There is another class of examples that I will not go into in detail here. This type has been discussed in the literature on Yiddish (Cable 2004; Landau 2004; Davis – Prince 1986) predicate clefts among others under the term **pseudo-infinitives**. These are infinitival forms that are "regularized" in the sense that they involve the mechanical adding of the infinitival suffix (*-ni* in Hungarian) to the root, without regard to irregular infinitival forms. Observe the Yiddish and the Hungarian examples below:

(i) **Veysn / *visn veyst** zi es. (Yiddish; from Cable 2004)
know-inf knows she it
'She KNOWS it.'

(ii) **Van-ni / *lenni** van pénzem (de nem elég).
is-inf / be-inf is money-mine (but not enough)
'It's true that I have some money but not enough.'

(iii) ?**Van-ni nincs** (de megpróbálok szerezni).
is-inf not-is (but I-try get-inf)
'We don't actually have any but I will try to get some.'

(iv) ?**Vol-ni vol-t** (de elfogyott).
was-inf was-past (but ran-out)
'It's true that there was some but it ran out.'

Although the last two examples are slightly marginal, the pseudo-infinitive of *be* occurs in (16b)-type sentences often—actually, the normal infinitive *lenni* of *be* is not possible in such existential constructions. This is an intriguing fact, pointing to an analysis of the *be* of existential sentences (which is most likely focused but in any case sentence-initial except for topics) that sets it apart structurally from the copula (which cannot occur in fronted predicates, and is probably a spell-out of T) as well as from the *be* inserted under V (which occurs in fronted predicates in its "normal" irregular form).

2. Proposal about the nature of dative case in Hungarian

Now let us return to the question of why and how fronted nominal and adjectival predicates end up with dative case. Since I hope to show that this emergence of dative is not exceptional at all, let me summarize the general occurrences of dative in Hungarian. It should be noted from the discussion below that (a) most datives in this language are predicative; (b) secondary predication in particular shows a similar distribution of infinitives and datives; and (c) it can therefore be claimed that when a predicate is "secondary" in the precise sense to be defined below, it surfaces as an infinitive if it is a verb, and as a dative if it is a noun or adjective.

(18) (a) Goals: Péter-nek adtam az összes pénzem.
Peter-dat I-gave the all money-my
'I gave all my money to Peter.'

(b) *Have*-sentences: Péter-nek rengeteg pénze van.
Peter-dat a-lot-of money-his is
'Peter has a whole lot of money.'

(c) Raised possessors:[15] Péter-nek elveszett a pénze.
Peter-dat got-lost the money-his
'Peter's money has been lost.'

(d) Beneficiaries: Péter-nek élek.
Peter-dat I-live
'I live for Peter.'

Although I will only concentrate on standard cases of secondary predication while making my argument, note that other occurrences of the dative in Hungarian also involve predication, at least on certain theories

[15] I am assuming (contra Tóth 2002; in the basic spirit of É. Kiss 2001) that dative "subjects" of inflected infinitives can be analyzed as possessors, and thus as predicates. An example of this construction is:

(i) Péter-nek men-ni-e kell.
Peter-dat go-inf-3sg must
'Peter has to go.'

For arguments for and against the possessor-analysis of these datives, see the above papers. In order to avoid overcomplicating this discussion, I do not take up this issue here.

and assumptions.[16] In particular, if we predict that dative case and an added layer of predication go hand in hand, this provides support for analyses of goals and possessors as predicates in their own right. Since this is not crucial to the point I wish to make here, I will use as illustration occurrences of the dative case that are unarguably predicates under any theoretical assumptions.

The most straightforward instances of dative predication are as follows:

(19) (a) Péter **okos-nak** / **zseni-nek** látszik.
Peter smart-dat / genius-dat seems
'Peter seems like a smart man/a genius.'

(b) Péter-t **okos-nak** / **zseni-nek** tartják.
Peter-acc smart-dat / genius-dat they-consider
'Peter is considered smart/a genius.'

(c) Péter **politikus-nak** készül.
Peter politician-dat prepares
'Peter is preparing (planning) to become a politician.'

(d) Péter-t **elnök-nek** választották.
Peter-acc president-dat they-elected
'Peter has been elected president.'

Examples (19a–d) show standard cases of secondary predication, meaning that the adjective or noun is obviously predicated of the DP (*Péter(t)*) that has become (depending on the selectional grid of the matrix verb) the main clause subject or object. The interesting fact about these Hungarian cases is that the secondary predicate itself occurs in dative case, and it occupies the preverbal position in the matrix clause. In effect, it has formed a complex predicate with the matrix verb, something like "seems to be a genius" or rather "genius-(to-be)-seems" — so its status is clear, at least from an interpretational point of view.

Now, the question still remains: where does the dative case come from? If we are to posit a local relationship between the subject and the secondary predicate in (19a–d), for example, the most standard way to achieve this is through a small clause — without a verb, naturally, since

[16] In particular, see Marcel den Dikken's well-known work on the possessive construction, arguing that the possessor should be analyzed as a predicate, rather than a subject, in these phrases. See for example den Dikken (1999) for discussion.

there is no copula in Hungarian where there is no tense to be spelled out. This makes it implausible, however, that the matrix verb should be the dative case-assigner: once the SC is formed, the subject must raise for case, but the predicate (which, by definition, does not need case) has no motivation for movement. And even if the motivation were simply structural (some form of clause-union or predicate-incorporation), this would not explain why case assignment is required. This makes natural the suggestion that dative case is assigned inside the small clause.

There is an alternation that, in my opinion, sheds light on this issue:

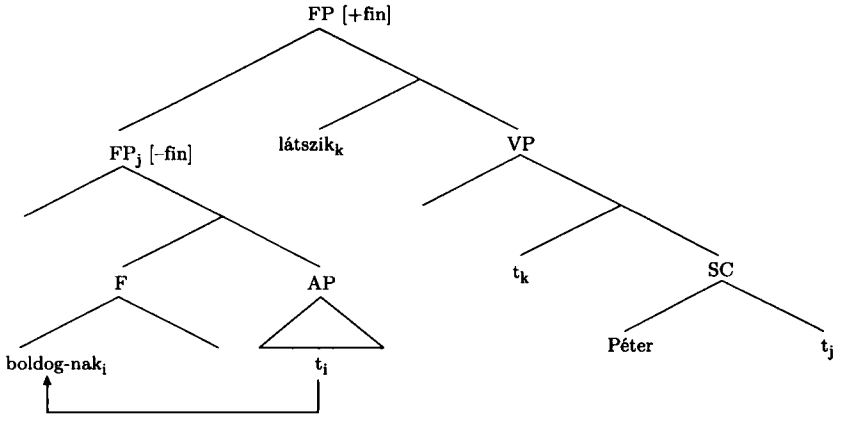
(20) (a) Péter **boldog-nak** látszik.
Peter happy-dat seems
'Peter seems happy.'

(b) Péter **örül-ni** látszik.
Peter be-happy-inf seems
'Peter seems to be happy.'

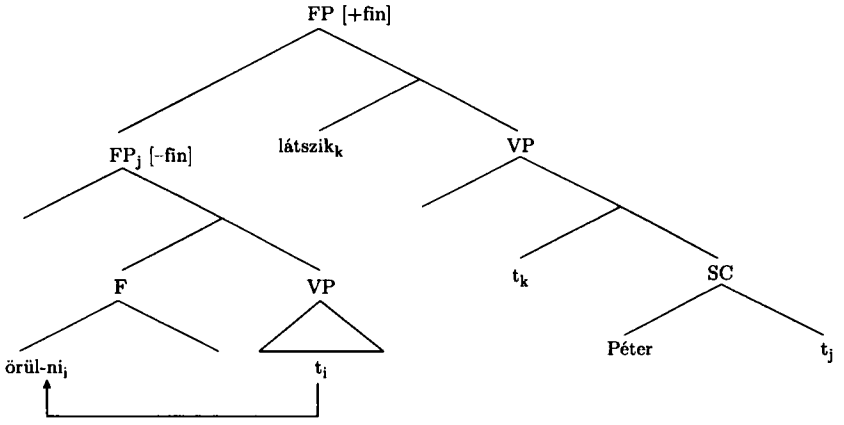
The examples in (20) are nearly synonymous, with the difference that the secondary predicate in (20a) is an adjective, and in (20b) it is a verb. Given that they both start out as predicates (of a small clause), it seems natural to assign them similar structures with a single difference: while (20a) has an adjectival predicate in the small clause, (20b) contains a verb. I suggest that, in both cases, the lexical predicate head-moves to F (the head of the functional projection dominating the lexical projection of a predicate), FP is later moved to the PM position of the matrix clause ([SpecFP] dominating the matrix VP), and the SC predicate is spelled out in this position—as dative in the adjectival case, and as infinitive in the verbal case. The derivation is shown below (without the subsequent movement of the SC subject to Spec,TP and later to Topic):[17]

[17] In the trees, I have added [+ fin] and [− fin] for ease of exposition; I am not implying that these two F heads are inherently different—but F[+ fin] is immediately dominated by T, while F[− fin] is not. Since "F" may very well be the locus of Aspect (among other things), it is plausible that it enters into some formal relationship with T.

(21) (a) Péter **boldog-nak** látszik.
Peter happy-dat seems



(b) Péter **örül-ni** látszik.
Peter be-happy-inf seems



Note that the structures above contain an empty Spec position left of the fronted predicate, which is in fact recursive, that is, the embedded predicate can also attract its own PM into its Spec, just like the matrix predicate attracted it. The fact that at least some modifier (in particular, the already mentioned predicate-modifier) moves along with the predicate to the matrix clause in these constructions is shown by examples like (22) below, where the verb is modified by a particle or a bare adverb, and (23), where the nominal or adjectival predicate has a modifier:

(22) (a) Péter **meg-hatód-ni** látszott.
Peter part-be-touched-inf seemed
'Peter seemed to be touched.'

(b) Péter **haza-indul-ni** készült.
Peter home-start-inf prepared
'Peter was preparing to leave for home.'

(23) (a) Péter-t **rendkívüli zseni-nek** tartják.
Peter-acc exceptional genius-dat they-consider
'Peter is considered to be an exceptional genius.'

(b) Péter **átlagos testalkatú-nak** tűnik.
Peter average built-dat appears
'Peter appears to have an average build.'

In (22)–(23), it is clear that the chunk moving to the PM position of the matrix clause (in bold above) must contain a phrasal layer since it contains a modifier. The assumptions about verbal predication—namely, that right outside the core VP there is a functional phrase (PredP or AspP in the literature), whose head houses the V and Spec houses the verb-modifier, a designated modifier of V that most affects its meaning and aspectual properties—extend naturally to predicates of other categories, as will be shown in section 4. The advantage of such uniform treatment is that no special provisions need be stated for secondary predication or small clauses in general: the predicate head (V, A or N) raises to F– in this case, non-finite FP is the maximal predicative projection in the lower clause; in other words, there is a full predicate (including modification and aspect) but there is no tense. Subsequently, this lower FP raises to form a complex predicate with the matrix verb in cases where the higher structure requires such movement—that is, where it later turns out that the finite predicate is a different one. Thus the structure of the relevant portion of (22b) becomes as below:

(22′) $[_{FP}\ [_{FP}$ haza $[_{F'}$ indulni$]]_i\ [_{F'}$ készült$_j\ [$VP ... $t_j\ [$SC ... $t_i]]]]$

Interim conclusion 2: Dative case is structural – it occurs when a nominal or adjectival predicate is spelled out (for any reason) in the head of a non-finite FP (that is, FP not directly dominated by Tense). This configuration obtains for one whenever a lower FP raises to a higher position and therefore the nominal/adjectival predicate raises no further than this position.

As I have implied above, a point pursued in this paper is that the occurrences of predicative dative case (or of the infinitival form of the verb) are structural in the sense of configurational. By this I mean the following: there are independent spell-out conditions dictating which copies in a movement chain are spelled out in any given structure. Normally, these conditions have the effect of forcing the spell-out of the highest copy of a lexical item, and this is so in all examples treated above. These conditions will, for example, take a structure like (21a) and state that the highest copy—sitting inside the raised FP—shall be pronounced. The **form** which this copy takes is dependent on morphology and various other lexical constraints— in this case, since the relevant copy is situated in the head of non-finite FP (= FP not directly dominated by T), it will be pronounced as dative, just as a verb in the same position in (21b) takes the form of the infinitive. Note that this is not a default form in any sense— no more is the infinitive a default verb than the dative is the default form of an adjective or noun. If anything, we expect secondary predicates to be more marked morphologically than their primary counterparts—and this is so. The point is (and this is strongly in accordance with Distributed Morphology) that **the spell-out of a verbal head in F (i.e., a predicative position without tense) is the infinitive**, just as **an A or N spelled out in the head of FP (again, a predicative position without tense) is the dative.**[18] As I plan to show, this point is nicely supported by predicate fronting.

As pointed out by many authors (on predicate fronting by Landau 2004), so-called "exceptional" constructions are often anything but exceptional. What they are is windows into aspects of the process of structure-building that are otherwise masked by surface syntax. In the case of predicate fronting, this idea manifests itself particularly clearly. While the clause itself is built up in the usual way, the non-finite version of FP (formed the same way in small clauses and in matrix clauses, since—barring look-ahead—these higher structural levels are not yet visible at this stage) also happens to be spelled out due to the requirements of Topic or Focus (i.e., the ban on null topics or foci). Through fronting,

[18] An issue for further research would be to see if such structural parallels between the dative and the infinitive can be extended naturally to other languages. One obvious candidate is English, where goals and infinitives are both introduced by the preposition *to* (which could easily be housed in F). Another place to look for such parallels might be languages like Japanese, where the verbal and adjectival predicates can appear with the same marking in certain configurations (e.g., predicate fronting; Tomoyuki Yabe, p.c.).

FP is removed from under the dominance of T and, as such, receives the same pronunciation as non-finite FP in (21) for example. Thus, rather than involving any form of "exceptional case-marking", predicate fronting supports an important point: that secondary predication has no special properties. To be more precise, if we posit that the lowest layers of structure are built up uniformly, based on simple lexical properties and syntactic principles, we must draw the conclusion that whatever morphological differences appear on the surface are due to higher structures and their requirements.

To illustrate, let us look again at the examples in the introduction, which I have claimed involve the same FP-formation, and a basic, neutral sentence:

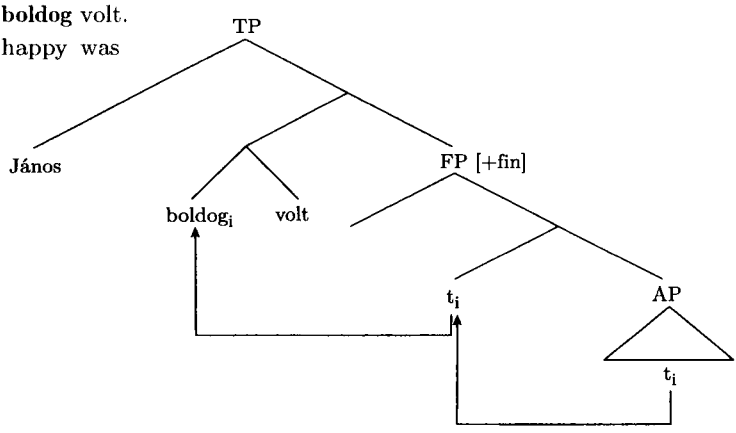
(24) (a) János **boldog** volt.
John happy was
'John was happy.'

(b) János **boldog-nak** látszik.
John happy-dat seems
'John seems happy.'

(c) **Gazdag-nak**´ (János) **gazdag** volt (de mégsem volt boldog).
rich-dat (John) rich was (but yet-not was happy)
'Rich he (John) was, but he still wasn't happy.'

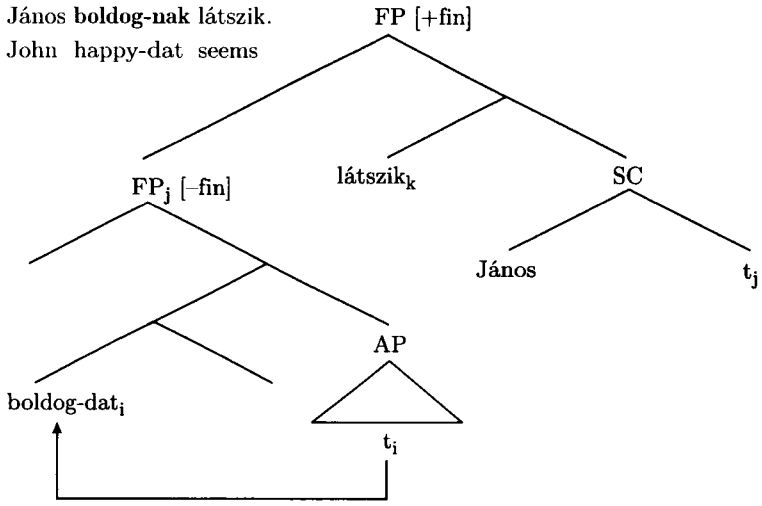
(24a) involves a simple sentence, therefore the highest occurrence of the adjective will move to T to support the inflection (spelled out as the copula), as shown below:

(24) (a′) János **boldog** volt.
John happy was



Let's now look at (24b). Unlike in (24a), here the head position of the lower, non-finite FP (which is not directly dominated by Tense) is the highest position reached by the adjectival predicate—since the finite predicate happens to be a different one, and therefore the adjective stays put in its first-derived position. Thus, the morphological configuration for dative spell-out obtains:

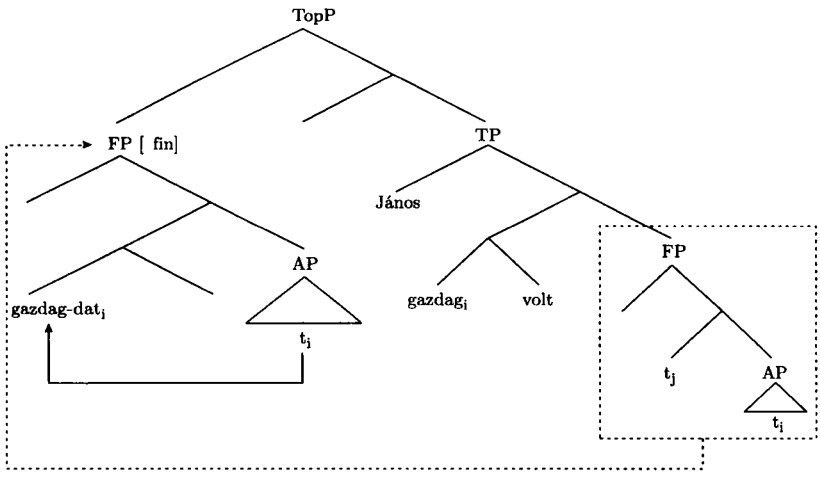
(24) (b′) János **boldog-nak** látszik.
John happy-dat seems



Finally, (24 c′) displays what is predicted for the structure of predicate fronting. In a sense this structure is a combination of (24a) and (24b): while the adjective moves out of its own projection to become the main predicate, its copy in the head of FP is also spelled out due to the fact that no null topics are allowed—therefore, when the chunk moved into Topic happens to be an FP undominated by Tense, its content must be pronounced as non-finite—that is, as an infinitive if the predicative head is a verb, and as dative if the head is an A or N.

Notice that the mere fact that an adjective or noun is moved to contrastive topic does not result in dative case assignment, observe (25).

(24) (c′) **Gazdag-nak´** (János) **gazdag** volt...
rich-dat (John) rich was



(25) (a) Boldog´ volt János de gazdag nem.
happy was John but rich not
'John HAS been happy (in his life) but not rich.'

(b) Boldog-nak´ boldog volt János de mégis aggódtunk érte.
happy-dat happy was John but still worried-we for-him
'Happy John was but we still worried about him.'

While (25a) involves movement of the AP into contrastive topic,[19] (25b) is the now-familiar FP-fronting. The difference is apparent from the interpretation: while in (25a) it is merely the property of happiness that is contrasted with other properties, in (25b) the state of being happy is contrasted with other related facts (here: the fact that we still worried about him). At least according to some speakers, AP-fronting such as in (25a) is only possible on an in-his-life interpretation, where the copula is in focus—presumably this higher movement of the tense head frees up the predicate itself for topicalization. Or rather, the fact that the copula constitutes the main assertion in (25a) degrades the AP to the status of modifier. But the point here is that the presence of dative in (25b) signals

[19] Consequently, this is a simple case of constituent topicalization, available in Hungarian for DP's most generally, but (as we see here) also for larger units. Whether any phrases larger than FP can be fronted is an open question.

that something else is at work—according to this paper, the something else is the fronting of a larger chunk, namely FP.

Crucially, the fact that the adjective is pronounced with dative and verbs come out as infinitives in this position provides support for the claim that it is this same FP that is involved in secondary predication. It also shows that secondary predication is a derivational concept, that predicates move through the same configuration in small clauses as they do in finite sentences—if this were not so, we would have no plausible explanation for why predicate fronting examples like (24c) should involve "secondary predicates". The simple intuition is that predication exists with different layers—thematic, aspectual, temporal—and these layers may coincide (as in simple sentences) or not (as in "secondary predication")—and may be teased apart thanks to "exceptional" phenomena such as predicate fronting.

3. Is there movement in predicate cleft constructions?

One question that is dealt with in detail in papers about predicate clefts (and about topicalization in general) is whether these constructions involve movement at all. Therefore, before I get into the question of **what** the fronted constituent is, I will briefly address the question of whether non-movement analyses are plausible or worth pursuing.

While it is quite obvious that the secondary predication cases discussed in the previous sections involve movement, the situation is much less clear when it comes to predicate clefts. There are two immediately obvious facts clouding the picture: firstly, there are two copies of the same element spelled out, and secondly, there is the well-documented mismatch between the two copies. As I have shown above, the appearance of dative case (and the infinitive ending) in Hungarian predicate clefts is not at all unexpected, so this factor is not decisive between a movement and a non-movement analysis. Leaving aside the question of multiple spell-out (and assuming that multiple pronunciation in itself is not an argument against movement), let me now turn to the question of whether any direct evidence can be found for fronting.

The standard tests for movement are also the ones cited in the predicate cleft literature as evidence for fronting.[20] The same island-effects can be reproduced for Hungarian:

[20] See the references at the end of this paper.

(26) Factive island:

(a) Bátor-nak azt hiszem / tudom, hogy bátor volt.
brave-dat that-acc I-believe I-know comp brave was
'As for brave, I believe/know he was. (But he still didn't become a good soldier.)'

(b) *Beteg-nek sajnálom, hogy beteg volt.
sick-dat I-regret comp sick was
'I do regret that he was sick. (But I'm still not sorry he didn't come to the party.)'

(27) Wh-island:

(a) Beteg-nek meg-mondom neked, hogy mikor volt beteg utoljára.
sick-dat I-tell you comp when was sick last
'As for being sick, I can tell you when he was last sick. (But how will that help you figure out if he has taken all his sick leave?)'

(b) ??Beteg-nek meg-kérdeztem, hogy mikor volt beteg utoljára.
sick-dat I-asked comp when was sick last
'As for being sick, I asked him when he was last sick. (But I forgot to ask whether he took sick leave that time.)'

Such tests standardly show that there is a movement relation between the base-copy and the fronted copy at some point in the derivation. However, as pointed out by Cable (2004) and Vicente (2005) on different grounds, they do not actually rule out the base-generation analysis. Since all of the examples are biclausal, technically an analysis is possible where the dative version of the predicate is generated on the left edge of the lower CP, and is subsequently moved to the topic position of the matrix clause. In such a scenario we would also expect to observe the island effects exemplified above (and in fact this analysis is argued for in Cable 2004).

However, as pointed out by Vicente (2005) in a paper on Spanish predicate clefts, monoclausal tests can also be constructed to detect movement on a local level. In particular, he observes that there are two dialects of Spanish differentiated precisely by their behavior with respect to such tests. The test used by Vicente to prove his point is the well-known coordination island test, shown below for Hungarian:

(28) *Szép-nek szép és okos volt (de...)
pretty-dat pretty and smart was (but...)
'As for being pretty, she was pretty and smart (but...)'

The coordination test shows local movement, and divides the dialects of Spanish into two groups: Dialect A, which patterns with Hungarian — where association of the topic with just one of the conjuncts is illicit, showing movement; and Dialect B, which should pattern (at least superficially) with Yiddish, on which Cable bases his analysis (although Cable does not discuss these tests) — where (28) is grammatical. As Vicente (*ibid.*) observes, the availability of sentences like (28) corresponds with whether or not so-called "genus-species" predicate clefts are possible. This odd class of predicate clefts is allowed in a subset of the languages or dialects that have predicate clefts at all (for example, in Yiddish), and look like the example given below:

(29) Essen fish est Max hekht.
eat-inf fish eats max pike
'As for eating fish, Max eats pike.'[21]

The obvious characteristic of these "genus-species" sentences is that the element in the topic position is a more general term for what appears as the predicate inside the clause. As Vicente argues (and as many authors have observed), it is difficult to analyze these cases as instances of movement. Rather than take their mere existence as counterevidence for movement, however, Vicente observes that the availability of "genus-species" predicate clefts correlates with the possibility of associating one conjunct of a coordinate structure with the topic. As expected, Dialect B above, which allows examples like (28), also has no problem with sentences like (29). It can be concluded, therefore, that there is a dividing line between two types of predicate clefting: one (as in Dialect A) involving movement of some sort, and one involving base-generation.

As predicted, Hungarian also does not allow "genus-species" sentences, just like Dialect A of Spanish, and as shown below:

(30) (a) *Hal-nak harcsa volt, amit fogott.
fish-dat catfish was what-acc he-caught
'As for fish, what he caught was a catfish.'

(b) *Szép-nek gyönyörű volt, nekem mégsem tetszett.
pretty-dat beautiful was me still-not appealed
'As for being pretty, she was beautiful, but I still didn't like her.'

[21] The example is taken from Cable (2004).

Therefore, without taking a stance on Dialect B-type predicate clefts, I will henceforth assume that Hungarian predicate clefts do in fact involve movement, and thus fall into the same class as Dialect A-type Spanish dialects and all other languages that do not allow "genus-species" topics.

4. What moves?

Thus far I have been using the label "FP" in a slightly vague manner, which I will attempt to make more explicit below. In particular, I will show what can and must move along with the predicate in predicate fronting, in what way these elements form a natural class, and how to go about ruling out the fronting of other complements and modifiers that do not appear in Topic. As before, I refer to the projection just outside VP/AP/NP as FP, although nothing hinges on this decision, and a shell-based approach would enable the same result.[22]

To begin, let me mention an issue that comes up often in the literature on predicate clefts—namely, whether the movement involved affects a phrase or a head. This question is a natural one to raise for languages where no modifiers can move along with the predicate (actually, most literature on these constructions only deals with verbs). In Hungarian, the situation is obviously not so simple, since at least particles and the so-called "predicate-modifiers" can (and, if they are present, must) front along with their predicate, making something like (31) ungrammatical:

(31) (a) *Men-ni haza-mentem (de már nem maradt időm pihenni).
go-inf home-I-went (but already not remained time-mine rest-inf)
'I did go home, but I didn't have any time left to rest.'

(b) *Testalkatú-nak átlagos testalkatú.
built-dat average built
'As for his build, he has an average build.'

However, when the predicate occurs most naturally without any modifier in a neutral sentence, it obviously also fronts without a modifier, so the question of head vs. phrasal movement may be relevant in these cases.

[22] That is, vP involving the movement of V to light v is pretty much the same thing, except perhaps for the locus of categorical information. However, FP is intended to be more general, as it allows us to treat all predicative constructions analogously.

In fact, it is argued in Landau (2004) for Hebrew that "when stranding its arguments, the fronted category [in predicate clefts] is a bare V (rather than a remnant VP)." As Landau's Hebrew facts show, an internal argument of the verb may be fronted along with it or stranded, as shown below:

(32) (a) [liknot et ha-praxim], hi kanta
buy-inf acc the-flowers she bought
'As for buying the flowers, she bought (them).'

(b) [liknot], hi kanta el ha-praxim
buy-inf she bought acc the-flowers
'As for buying, she bought the flowers.'

Simplifying Landau's analysis somewhat, his idea is that while an example like (32a) arises through VP-fronting and subsequent deletion of the subject, (32b) comes about via fronting of the V head to topic (so no remnant movement or deletion is required).

Notice that the issue arises for Hungarian in a slightly different way. There are three varieties of the same example to consider, all with distinct syntactic and semantic properties:

(33) (a) [Virágot venni] virágot vett...
flower-acc buy-inf flower-acc bought
'As for buying flowers, that's what he did (but I don't know where he did that).'

(b) [Venni] vett virágot...
buy-inf bought flower-acc
'Buy he did some flowers (but not nearly enough for the decoration)'

(c) [Venni] VIRÁGOT vett...
buy-inf flower-acc bought
'What he bought was flowers (but what he sold was fruits).'

(d) *[Venni virágot] (virágot) vett (virágot)...

As is apparent from the final, ungrammatical example (and as shown extensively above), post-head modifiers are ungrammatical in the fronted phrase. At the same time, whether or not the pre-head modifier will front along with its predicate depends on what actually forms the basis of the contrast. In (33a), the contrast is between the situation or act of flower-buying and other possible scenarios. For example, the previous speaker may have asked "Didn't he buy flowers (=go flower-buying) earlier?"

where the reply contains agreement with the whole predicate "flower-bought". On the other hand, (33b–c) contain examples where the basis of the contrast is simply the verb ("buying"); in the first case, the previous utterance may have been something like "Are you telling me that he didn't buy any flowers?" while in the second case "Didn't he sell flowers? I mean, I know that he bought flowers..." The crucial point is that it would appear from the interpretation that examples (33b–c) may be analyzed as topicalizing only the verbal head.

However, the situation is not as clear as that. Firstly, it must be noted that the difference between examples (33a) and (33b–c) lies not only in the content of the fronted element, but also in the structure of the base sentence. While the direct object is clearly a verb-modifier in (33a)—apparent from its neutral position left-adjacent to the inflected verb—, this is not the case in (33b–c). In (33b), the inflected verb is in focus—shown by intonation and by the stranding of its accusative argument;[23] in (33c), the direct object is in focus. Without dwelling too much on the Hungarian-specific details of where exactly the base sentences of the examples in (33) differ structurally, it can be concluded that **whenever the predicate-modifier position (here: SpecFP) is filled in the base sentence, the element filling that position must front along with the predicate.**[24] Thus, the fact that the predicate head appears to be alone in Topic in (33b–c) may well indicate that the relevant modifier position is simply empty or filled with a silent element.[25] Therefore, I

[23] This is the normal state of affairs in Hungarian, where "in-his-life" or "accomplishment" readings—loosely: the English present perfect—are usually obtained by focusing the verb, as in:

(i) Vettem kenyeret. (cf. Kenyeret vettem. 'I bought bread.')
I-bought bread
'I have bought bread.'

[24] This correlation makes a head-movement analysis very implausible for Hungarian: There would need to be some restriction stating that phrasal movement is required whenever the PM position is filled, and head-movement is required otherwise. (Note that nothing other than the PM is ever allowed to surface in the fronted phrase.)

[25] In fact, it is likely that in the case of verbal focus (such as (33b)), there is some silent adverbial element filling the PM position—as shown by the impossibility of having an overt PM in such sentences:

(i) *(Meg-látni) meg-láttam Párizst életemben.
(part-see-inf) part-saw paris-acc in-my-life
'(As for seeing it,) I have seen Paris in my life.'

will not take the occurrence of infinitives without complements in Topic as an argument for head-movement, and continue to analyze these as instances of phrasal fronting where the phrase happens to contain only one overt element.[26]

A related fact observed in the literature is the difference between the fronting possibilities of copular *be* and verbal *be*[27] (i.e., the *be* that is modified by an adverb). In Hebrew, for example, copular *be* cannot front by itself; this *be* can only be topicalized along with its complement — an NP or AP. However, verbal *be* can appear alone in topic. At the same time, it's not clear how such a restriction should be formulated. Landau (2004) is led to conclude that (since auxiliaries also cannot front by themselves) "semantic richness" is the decisive factor here but the difference strikes me as more of a structural one: verbal *be* is inserted under V, therefore we expect it to behave like other verbs. In particular, it (just like verbs in general) can be fronted alone as long as its complement is focused; observe an example from Hungarian:

(34) [Lenni] MELLETTEM volt, de az agya láthatóan máshol járt.
be-inf next-to-me was but the brain obviously elsewhere went
'He was actually next to me, but his mind was obviously somewhere else.'

If no such focusing takes place, verbal *be* takes its modifier along to its fronted position:

(35) [Otthon len-ni] otthon volt (de nem nyitott ajtót)
home be-inf home was (but not opened door-acc)
'As for being at home, she was, but she wouldn't open the door.'

[26] As for the Hebrew examples cited above, it should be observed that there is also an interpretational difference between the two — namely, (32b) is similar in interpretation to (33c), involving focus on the direct object. This is noted in Landau's paper but he dismisses it as a possible motivation for the difference between (32a–b) based on internal properties of Hebrew grammar (facts related to the position of negation and the form of pronouns occurring in this position). I will not comment on his examples here in particular, merely noting that a very similar problem in Hungarian can be resolved by taking into consideration the (visible) focus movement that precedes the fronting of the predicative phrase. Whether this line of analysis can be stretched to cover the Hebrew cases is unclear at best at this point.

[27] Landau does not use this terminology; I have added it for clarification. Note that by using these shorthand terms I do not wish to imply that the two types of *be* are separate lexical entries. On the contrary, I am assuming that the same (semantically impoverished) set of features can be inserted under different nodes to different structural effect.

On the other hand, copular *be* is inserted (at least in Hungarian) higher than FP (probably under T), and therefore it is not present when FP-fronting takes place:[28]

(36) Büszké-nek (*volt / *lenni) büszke volt.
proud-dat was to-be proud was
'As for being proud, he was.'

Since all of these examples point in a single direction (namely, that the phrase affected by Hungarian predicate fronting always displays the same properties), I will disregard the possibility of head-movement to Topic, and concentrate on the identity of the fronted phrase.[29]

In the next short section I will attempt to characterize the projection I have been referring to as "FP" and provide an explanation for the special status of the modifier I have labeled "PM". In doing so, I hope to motivate a view of Hungarian syntax I have defended elsewhere and in different contexts: that the special status of the PM is due to the fact that it is merged lowest in the course of the derivation (in a predicative position)

[28] In Hebrew, the situation is obviously different, given that the copula can surface in the fronted constituent, while the fronting of the copula alone is ungrammatical, as seen below (Landau's example):

(i) Lihyot zamin, Gil lo tamid haya.
to-be available Gil not always was
'As for being available, Gil wasn't always.'

(ii) *Lihyot, Gil lo tamid haya zamin.

It seems to me that a reasonable explanation for this contrast cannot be semantic, since BE is semantically empty by definition. Nor is it easy to formulate a plausible rule restricting head-topicalization to verbal heads. Without attempting any serious analysis of the Hebrew facts, I would like to (tentatively) suggest that—since it does appear in the fronted phrase—the Hebrew copula is inserted lower than the Hungarian one, perhaps under F, and the non-verbal predicate head does not raise to it but remains inside its own phrase. (It appears after the copula in the fronted phrase, and is deleted from its base position, on a par with other complements.) In this case, the problem is re-formulated as follows: Why is it that the AP or NP cannot be focused or otherwise scrambled out of its base position (dominated by FP)? In a way, this is a similar situation to that of idioms: verbs taking idiomatic complements also cannot be fronted alone. Whether this would involve semantic incorporation or actual (covert) movement is an issue whose resolution is outside the scope of this paper.

[29] For further arguments against head-movement in predicate clefts, see Abels (2001). For an alternative view, doing away with the distinction between head and phrasal movement altogether, see Nunes (1999).

—and ends up in (SpecFP) in the surface syntax of a neutral sentence. This line of analysis, in addition to making a lot of interesting predictions possible in other realms of syntax, also obviates the need for complicated definitions of the modifier of the relevant predicative phrase.[30]

It has long been known about the syntax of Hungarian that there is something special about the position left-adjacent to the verb in a neutral sentence (involving no focus or negation, which result in further verb movement, obscuring the lower positions). Primarily, the relationship between this (traditionally: VM = Spec,AspP or Spec,PredP) position and aspect has been investigated at length, the broad generalization being that the content of this position plays a crucial role in determining the aspectual and aktionsart properties of the sentence. In essence, the VM (= PM) forms the main predicate in the sentence. This is precisely the factor that obscures its status—see footnote 9 for some comments.

In sentences containing more than one clause, it can happen that the PM of the lower clause raises to the matrix predicate, if the latter has no modifier of its own:[31]

(37) El akartam napolni a problémát.
part wanted-I postpone the problem-acc
'I wanted to postpone the problem.'

The particle *el* belongs to the lower verb (cf. *Elnapoltam a problémát.* 'I postponed the problem.'), but it raises to the matrix clause to form a complex between the two predicates (= "postpone-want" or similar, although the lower infinitive cannot surface in the PM position for independent reasons). Now we come to a case mentioned above, which bears directly on the identity of the phrase fronted:

[30] This restriction sets the lower limit for the content of the fronted phrase but not the upper one: there may well be adjuncts (or even selected complements) inside the relevant phrase's base copy. These phrases do not have the special status of PM and therefore remain post-head. For simplicity's sake, I will assume in this paper that post-head material is deleted in Topic/Focus on identity with the lower copy, following well-established principles on the economy of pronunciation. In a phase-based approach it is very likely that FP will turn out to be a phase, in which case its domain has long been spelled out by the time topicalization occurs. Meanwhile its Spec and Head could be spelled out once more in a higher position.

[31] On the particulars of this, see É. Kiss (2002).

(38) El napolni el akartam (*napolni) (de nem hagyták).[32]
part postpone-inf part wanted-I postpone-inf (but they didn't let me)
'I did want to postpone it but they didn't let me.'

It is clear from the example that, whatever the phrase fronted, it must be quite low in the structure—since on higher surface levels the combination *el-napolni* does not exist. This means that the particle and its predicate must have combined at some lower level, from there the particle has raised to become the PM of the matrix predicate, and the lower FP was subsequently fronted. Multiple pronunciation obtains because the PM is needed in the topicalized phrase for interpretation, while it is obligatory before the matrix predicate because of structural reasons, making both alternatives below ungrammatical:

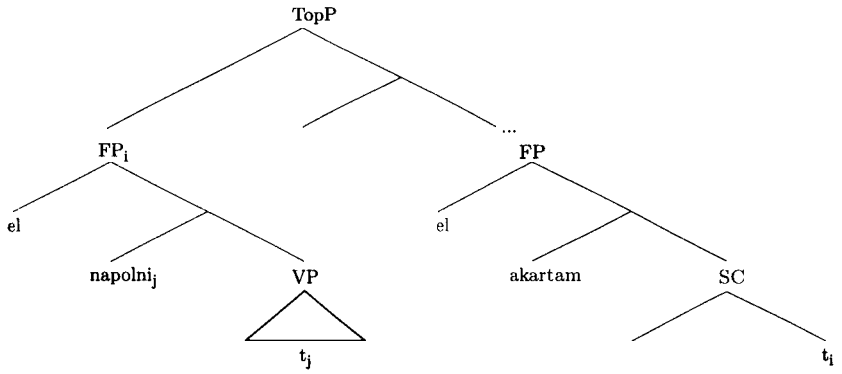
(39) (a) *Napolni el akartam.
(b) *El-napolni akartam. (cf. *Akartam el-napolni.)

Therefore, the structure of (38) is as given in (40).

While it is quite intuitive, as well as widely documented in the literature on Hungarian syntax, that verbs are accompanied by a modifier bearing a certain designated status, the same is not nearly so obvious when it comes to non-verbal predicates. There are two factors obscuring the parallel between the two types of predication. I will discuss them briefly in turn.

[32] This construction is very similar to one noted in Russian by Abels (2001).
(i) Čitat' (-to) on budet (*čitat').
read-inf to he will read-inf
'As for reading it, he will.'
Abels offers an explanation that captures the same idea outlined here, but in different words: doubling only happens when "the Base Line Sentence has only one exponent of both lexical content and tense information," which does not hold in complex verb forms. I would rephrase this as follows: "When the main verb is not the carrier of tense information (due to the presence of an auxiliary), it does not raise out of FP; therefore, when FP is fronted, it moves along with it, and is deleted on precise identity with the topicalized version." Anyway, this is what happens in Hungarian biclausal examples like (39) as well as analytic tenses exactly mirroring the Russian as in:
(ii) El-olvasni el fogom (*olvasni).
part-read-inf part will-I read-inf
'As for reading it, I will.'

(40) = (38)



The first issue is related to the status of the nominal and adjectival predicates themselves, namely, that often they appear as modifiers to the copula. However, it is obvious that the level at which the complex of the predicate and the copula is formed is higher than the predicative phrase affected by topicalization or secondary predication, shown by the unavailability of the copula in these fronted phrases:

(41) (a) Gazdag-nak (*lenni) gazdag volt.
rich-dat be-inf rich was
'As for being rich, he was.'

(b) Gazdag-nak (*lenni) tűnt.
rich-dat be-inf appeared
'He appeared rich.'

So it is most natural to assume that at the relevant level the adjective (or noun) **is** the predicate, and it is not a complement or modifier to the copula, which is inserted higher into the structure.

A more complicated matter, however, is that non-verbal predicates can themselves have modifiers of various sorts, as mentioned above, but the order of these does not appear to be as fixed as the order of modifiers to the verb — at least at first glance:

(42) (a) [A fiai-ra büszke apa] végig-tapsolta a meccset.
the sons-his-on proud father through-clapped the match-acc
'The father proud of his sons clapped all through the match.'

(b) *[A fiai-ra büszke] volt.
the sons-his-on proud was
'He was proud of his sons.'

(c) Büszke volt a fiai-ra.
proud was the sons-his-on
'He was proud of his sons.'

The issue is that while the adjective and its complement can evidently form a constituent in certain contexts (in an attributive position, as evidenced by (42a), they cannot form a complex predicate (shown by (42b)). This is loosely related to the observation that referential elements are generally not allowed as PM's—this position is restricted to predicative elements, which, by definition, cannot be referential. Thus, similarly to the situation with verbal predicates, certain modifiers of an adjective are most natural before the head, can appear along with the adjective in the position before the copula, and can front with the adjective to Topic:

(43) (a) [Közepesen magas] volt.
medium tall was
'He was medium height.'

(b) [Közepesen magas-nak] közepesen magas volt.
medium tall-dat medium tall was
'As for being medium height, he was.'

Without going into too much detail of this question, which requires a lot of further investigation, I will suggest that similar modification structures exist in verbal and non-verbal predication, which is natural if the relevant structural level (our FP) is the same, regardless of whether it dominates a VP, AP or NP.

Further support for the status of modifiers such as in (43) comes from answering patterns. In Hungarian, it is possible to give an affirmative answer to a yes-no question by uttering the PM (i.e., the main assertion) from the question:

(44) Q: El-napoltad a problémát?
away-postponed-you the problem-acc
'Did you postpone the problem?'
A: El.

This test can be used to distinguish between PM's and topics (which can also occur before the predicate, and although they cannot form a single intonational phrase with it, this is sometimes not so easy to tell). Now observe the distinct behavior of (42)- and (43)-type modifiers:

(45) Q: A fiaira büszke volt?
the sons-his-on proud was
'Was he proud of his sons?'
A: Büszke. / *A fiaira. / *Volt.

The pattern above shows that the main predicate in this sentence is *büszke*, which means that *a fiaira* must be a topic. (It can also stay low in its base position, of course.) Now compare:

(46) Q: Közepesen magas volt?
medium tall was
'Was he medium height?'
A: Közepesen. / *Magas. / *Volt.

The test indicates that the modifier *közepesen* constitutes the main assertion in the sentence, and must therefore be inside the predicative phrase. This, as mentioned above, correlates with the fact that it can (and must) front along with the adjective to TopP (see 43b).

Similar restrictions apply to nominal predicates, as shown below:

(47) (a) [Szigorú tanár-nak] szigorú tanár volt.
strict teacher-dat strict teacher was
'He was in fact a strict teacher.'
(b) Q: Szigorú tanár volt?
A: Szigorú. / *Tanár.
(c) *[Hivatalosan tanár-nak] hivatalosan tanár volt.
officially teacher-dat officially teacher was
Intended meaning: 'He was in fact an official teacher.'
(d) Q: Hivatalosan tanár volt?
A: *Hivatalosan. / Tanár.

Based on this drafty review of the relevant facts, I conclude that the analysis of the fronting of non-verbal predicates as FP fronting is well-founded, not only by analogy with the situation with verbal predicates

but also on independent grounds. This is desirable also because it provides a way to differentiate instances of FP-fronting on one hand and AP- or NP-fronting on the other, such as below:

(48) (a) Boldog´ volt János de gazdag nem.
happy was John but rich not
'John HAS been happy (in his life) but not rich.'

(b) Boldog-nak´ boldog volt János de mégis aggódtunk érte.
happy-dat happy was John but still worried-we for-him
'Happy John was but we still worried about him.'

(c) Tanár´ volt Judit de titkárnő nem.
teacher was judit but secretary not
'Teacher Judit has been, but secretary she hasn't.'

(d) Tanár-nak´ tanár volt Judit de csapnivaló.
teacher-dat teacher was judit but really bad
'Judit was in fact a teacher but she was pretty bad at it.'

Under this analysis, the difference comes from the size of the chunk fronted: AP in (48a), NP in (48c), and FP in (48b) and (48d). This correlates clearly with the availability of modification in the two cases. As shown above, there are looser restrictions on modification inside AP (42a) (which may allow phrase-internal topicalization) than in SpecFP. Observe:

(49) (a) Volt már életében [$_{AP}$ (a fiaira) büszke (a fiaira)]...
was already in-his-life the his-sons-on proud the his-sons-on
'He has been proud of his sons (in his life).'

(b) [$_{AP}$ (A fiaira) büszke (a fiaira)] volt már életében...
the his-sons-on proud the his-sons-on was already in-his-life
'As for being proud of his sons, he has been (in his life).'

(c) [(*A fiaira) büszke (*a fiaira)] volt a fiaira.
the his-sons-on proud the his-sons-on was the his-sons-on
'He was proud of his sons.'

(d) [$_{FP}$ (*A fiaira) büszké-nek (*a fiaira)] büszke volt a fiaira.
the his-sons-on proud-dat the his-sons-on proud was the his-sons-on
'As for being proud, he WAS of his sons.'

Therefore, I conclude that the status of the fronted predicate is (a) phrasal, and (b) distinct from the AP/NP containing the base copy of the predicate. I have labeled this phrase as FP.

5. Possible alternative analyses and comments on these

Before concluding and summarizing what I have said so far, I want to briefly mention some possible alternative analyses of the Hungarian facts, and counterarguments to these. Some of these alternatives have been proposed in the literature, although not all—due to the fact that this construction in Hungarian has not received much attention in the past. Therefore, I have tried to anticipate theoretically possible alternatives to my analysis. All of the analyses below—in addition to having inherent problems—share the property that they lack any kind of connection between the fact that dative case in predicate fronting occurs on predicates and that dative case is a normal way to mark secondary predicates in Hungarian.

5.1. Alternative Analysis A: Dative case comes from the topic position (structural)

This is clearly not the case, at least not in the strict sense—given that many kinds of elements can occur in that position without dative. One might argue, however, that predicates are the only elements that can be topicalized without already having received case—therefore the dative case associated with the position can only surface on them. This is again not true; observe an example similar to those already cited above:

(50) [Boldog] sokszor volt János (de gazdag nem).
happy many-times was John (but rich not)
'Happy John has been many times, but rich he has not been.'

When the AP, rather than the whole FP, is fronted—resulting in a contrast between two properties, rather than two predicates—no case occurs on the fronted predicate.

Also, the fronted predicative phrase can occur in focus rather than contrastive topic, as mentioned earlier as well, indicating that the dative case is not tied to a single structural position:

(51) [Csak boldog-nak] nem boldog (amúgy jól van).
only happy-dat not happy (otherwise well is)
'It's only that he is not happy, otherwise he is fine.'

Based on these examples, I believe that the view that dative case in these constructions should be structurally tied to the surface position of the predicate is not tenable. Rather, under my analysis, the dative occurs as a result of a low position of the predicate also being the spell-out position, due to higher levels of structure.

5.2. Alternative Analysis B: Dative case is inherent case, associated with the particular meaning 'as for the criterion of; from the aspect of...'

In Hungarian, there are some compelling arguments for this view, supported by cases such as:

(52) (a) Nekem / Péternek mindegy, mehetsz, ahova akarsz.
I-dat Peter-dat all-the-same you-can-go where you-want
'It's all the same to me/to Peter, you can go where you want.'

(b) Neked / ?Péter-nek ez kék?![33]
you-dat Peter-dat this blue
'This is blue in your/Peter's opinion?!'

These datives are always optional, and carry a very specific meaning: 'in the opinion of'. In this regard, they show a very close connection to the so-called 'ethical dative', as in:

(53) Már megint összetöri magát nekem / *Péternek.
yet again breaks self-acc I-dat Peter-dat
'He will get hurt "on me"/"on Peter" again.'

At the same time (in contrast with the predicate fronting datives), the types of dative in (52)–(53) are highly constrained. The ethical dative is always pronominal and clause-bound, while the opinion-datives in (52a–b) are also only possible with the types of predicates that allow the expression of an 'opinion' reading — compare:

(54) (a) *Nekem / *Péter-nek háromemeletes a ház.
I-dat Peter-dat three-storey-tall the house
Intended reading: 'In my/Peter's opinion, the house is three stories tall.'

[33] Thanks to Edith Kádár (p.c.) for calling my attention to this example.

(b) Háromemeletes-nek háromemeletes, de még így is a legalacsonyabb
three-storey-tall-dat three-storey-tall but still so also the smallest
az utcában.
the street-in
'It's true that it is three storeys tall but it's still the smallest in the street.'

I think it is clear from these examples that these two meanings of datives are not related. Nevertheless, the structural position of the ethical and opinion datives may also be syntactically fixed — this is something for further research.

5.3. Alternative Analysis C: Dative case is a default case in Hungarian (Burányi 2003)

Besides the fact that "default case" is a strange theoretical concept to begin with, there is no reason to assume that dative should be a default. Some of the cases cited elsewhere in the literature to support the idea of default dative — e.g., sentences like *It's me* in English — also involve predication; in fact, the 'dative' pronoun is a predicate. Nevertheless, to the extent that such a thing as 'default case' does exist, this does not explain its occurrence in the predicate fronting examples discussed here. Predicates do not need case, and so they don't need default case either.

Note that on the analysis advanced in this paper the same question can be turned around: if predicates do not need case, yet they bear dative case in a lot of examples, why and how do non-predicates receive dative? In other words, is dative case 'structural' in the sense that nominative and accusative are? The most desirable answer and an answer that I hope to investigate in upcoming work is naturally that dative always occurs on predicates.[34]

[34] Aside from theoretical objections, some of the data cited by Burányi in his paper is also questionable. Most importantly, he claims that verb modifiers that are not particles cannot be fronted along with the verb, making the example below ungrammatical:

(i) Kenyer-et ven-ni kenyer-et vett.
bread-acc buy-inf bread-acc bought
'As for bread-buying, that's what he did.'

Since according to my own judgment, as well as the opinion of about a dozen informants, the sentence above is well-formed, I can only assume that Burányi was working on a different dialect of Hungarian. Therefore, I believe it is pointless to

6. Summary and conclusions

In this paper I have presented data and arguments to support a uniform structure of FP dominating the lexical projection of the predicate—VP, NP or AP. I have argued that the internal properties of this phrase hold the key to various questions—among them the type of modification allowed in positions where FP is spelled out. I have shown that the internal modification patterns of FP and NP/AP are visibly different in Hungarian. I have also claimed that the dative and infinitive forms showing up in fronted predicates are a reflex of the F head; more precisely, the spell-out of V in F is the infinitive, and of N or A in F is the dative. As a side-issue I have suggested that the semantically empty features that constitute *be* can be inserted under the verbal node in Hungarian, yielding an element that behaves like a verb in the relevant respects, or these features can surface under T, which is the traditional "copula". This distinction explains the fact that the "copula" does not occur in the topicalized predicative phrase—since it is inserted higher than FP, its appearance would be unexpected.

One of the interesting characteristics of the internal structure of FP in Hungarian is the piling up of lower predicative elements in the specifier of the matrix predicate. In an "upward" derivation we witness each lower predicate to pick up and move into FP, which in turn moves up to become the Predicate Modifier of the matrix predicate and so on. According to this observation, we can expect various elements to start out from a predicative position, which is strongly in line with a wide variety of current research.[35] More strongly, a strict correlation between the surface position (Spec,FP) and the base position (inside a lexical projection dominated by FP) can be seen. This is intuitively appealing, since the idea that every sentence should only have one "main" predicate (however complex) is not a strange one. Hungarian may be a language where the various layers of predication unite to a large extent in overt syntax. Naturally, a question that I have left open—the exact nature of FP (=AspP, PredP, FocP or something entirely different)—will play a crucial role in formalizing this intuition.

---

go into the details of his analysis—given that one of the cornerstones of my paper has been that verb-modifiers can and always must front with their predicate.

[35] See for example den Dikken (forthcoming); Larson (2004); on Hungarian: É. Kiss (2002) for a predicative analysis of verb-modifiers, and so on.

A related issue is the status of datives in this picture. I have suggested that this research is aimed at supporting a more general line of thought, namely, that datives are predicative elements. This is a very appealing prospect,[36] and I believe that the properties of datives in Hungarian add significantly to the plausibility of this programmatic claim. In other terms, the idea that "case" should be correlated with features already necessary in syntax (i.e., the connection between nominative case and tense, accusative case and aspect) has been around for a long time. As for dative case, it has received slightly less attention, although the relationship between dative and the complexity of events (here: the complexity of predication) is discussed at length for example in Svenonius (2002). In my opinion, the implications of research along these lines are extremely promising and far-reaching.

In upcoming research, I hope to untangle many of the above mentioned matters.

References

Abels, Klaus 2001. The predicate cleft construction in Russian. In: Steven Franks – Tracy King – Michael Yardoff (eds): Formal approaches to Slavic linguistics 9: The Bloomington Meeting 2000, 1–18. Michigan Slavic Publications, Ann Arbor.

Bowers, John 2001. Predication. In: Mark Baltin – Chris Collins (eds): The handbook of contemporary syntactic theory, 299–333. Blackwell, Malden MA & Oxford.

Burányi, Péter 2003. Kettős predikálás: ismételt predikátumok kontrasztív topik pozícióban [Double predication: Repeated predicates in contrastive topic position]. In: Zsuzsanna Gécseg (ed.): LINGDOK 2. Nyelvészdoktoranduszok dolgozatai [Papers by doctorate students of lingustics], 48–70. Szegedi Tudományegyetem, Nyelvtudományi Doktori Iskola, Szeged.

Cable, Seth 2004. Predicate clefts and base-generation: Evidence from Yiddish and Brazilian Portuguese. Paper presented at the ECO5 Syntax Workshop, University of Maryland.

Davis, Lori J. – Ellen F. Prince 1986. Yiddish verb-topicalization and the notion "Lexical Integrity". In: Anne M. Farley – Peter T. Farley – Karl-Erik McCullough (eds): Papers from the 22nd Regional Meeting, 90–7. Chicago Linguistic Society, Chicago.

Dikken, Marcel den 1999. On the structural representation of possession and agreement: The case of (anti-)agreement in Hungarian possessed nominal phrases. In: István Kenesei (ed.): Crossing boundaries: Advances in the theory of Central and Eastern European languages, 137–78. John Benjamins, Amsterdam.

[36] I refer the reader to den Dikken's extensive work on possession and related matters.

Dikken, Marcel den forthcoming. Relators and linkers: The syntax of predication, predicate inversion, and the copula. MIT Press, Cambridge MA.

É. Kiss, Katalin 2001. A főnévi igenév személyragozásának kérdéseiről [On inflected infinitives]. In: Marianne Bakró-Nagy – Zoltán Bánréti – Katalin É. Kiss (eds): Újabb talulmányok a strukturális magyar nyelvtan és nyelvtörténet köréből [New studies on the structural grammar of Hungarian and the history of language], 249–73. Osiris Könyvkiadó, Budapest.

É. Kiss, Katalin 2002. The syntax of Hungarian. Cambridge University Press, Cambridge.

Komlósy, András 1992. Régensek és vonzatok [Valence and government]. In: Ferenc Kiefer (ed.): Strukturális magyar nyelvtan 1. Mondattan [A structural grammar of Hungarian, Vol. 1: Syntax], 299–527. Akadémiai Kiadó, Budapest.

Landau, Idan 2004. Constraints on partial VP fronting. Manuscript. Ben Gurion University.

Landau, Idan 2006. Chain resolution in Hebrew V(P)-fronting. In: Syntax 9:32–66.

Larson, Richard 2004. Sentence-final adverbs and scope. In: Matthew Wolf – Keir Moulton (eds): Proceedings of NELS 34, 23–43. GLSA, Amherst.

McCoy, Svetlana 2002. Semantic and discourse properties of colloquial Russian construction of the form "X-to X, a...". In: Glossos 3. (http://seelrc-web.aas.duke.edu/glossos/)

Nunes, Jairo 1999. Linearization of chains and phonetic realization of chain links. In: Samuel David Epstein – Norbert Hornstein (eds): Working minimalism, 217–49. MIT Press, Cambridge MA.

Pesetsky, David 1995. Zero syntax: Experiencers and cascades. MIT Press, Cambridge MA.

Svenonius, Peter 2002. Case is uninterpretable aspect. In: Angeliek van Hout – Henriette de Swart – Henk Verkuyl (eds): Proceedings of the Perspectives on Aspect Conference. University of Utrecht, Utrecht. (http://www-uilots.let.uu.nl/conferences)

Tóth, Ildikó 2002. Can the Hungarian infinitive be possessed? In: István Kenesei – Péter Siptár (eds): Approaches to Hungarian 8 (Papers from the Budapest conference), 133–60. Akadémiai Kiadó, Budapest.

Vicente, Luis 2005. Towards a unified theory of movement: An argument from Spanish predicate clefts. In: Leiden Papers in Linguistics 2.3:43 67.

Acta Linguistica Hungarica, Vol. 53 (3), pp. 333–339 (2006)
DOI: 10.1556/ALing.53.2006.3.4

BOOK REVIEW

László Hunyadi: Hungarian sentence prosody and Universal Grammar – On the phonology syntax interface. Peter Lang, Frankfurt am Main, 2002, 328 pp.

1. Introduction

The aim of the book is to show that sentential prosody, in particular intonation and stress, are vital tools available to the hearer in his/her task to uncover the logical proposition intended by the speaker. The author seeks to uncover the mechanisms of "the direct relation between LF and PF" (15). His main claim is that operator–scope relations are marked prosodically: in particular, languages may put main stress on either the operator or its scope. Hunyadi applies this to many attested word order variations and potential scope-readings of complex Hungarian sentences involving more than one operator. Ultimately, Hunyadi's goal is to derive the well-known ordering of constituents in the Hungarian left periphery (i.e., Topics–Quantifier–Focus–Verb). He also extends the cross-linguistic coverage of the theory to Finnish, Japanese, Hebrew and English.

The connections between semantics/pragmatics and prosody form the topic of much contemporary research, including, for instance, Reinhart's (1995) work on the prosodic nature of focus-marking in English, Büring's (1997) work on topic and focus in German, Zubizarreta's (1998) work on focus in Spanish, or Frascarelli's (2000) work on focus in Italian. An important aspect of Hunyadi's book that marks it out compared to its peers is that the data that forms the basis of the proposed prosodic theory is the result of phonetic experiments. This is very important, given the often controversial nature of judgements about stress placement and relative strength of stresses. Hunyadi's main claim that operator–scope relations may directly effect the prosody of an utterance is highly innovative. The variety of data that he encompasses is exceptionally large, involving the interactions between many different Hungarian operators such as *csak* 'only', the FOCUS operator, diverse types of topics, negation and various

1216–8076/$ 20.00

types of quantifiers. Even though this is an attractive feature of the book, it seems to me that sometimes the net is cast too wide and the detailed analysis of the wealth of facts is less precise than one would have liked. All in all, I believe that Hunyadi's *Hungarian Sentence Prosody and Universal Grammar—On the Phonology-Syntax Interface* should be read by anybody interested in the syntax–phonology interface, the interaction between discourse and prosody or the syntax of the "left periphery".

There are ten chapters altogether, and two appendices. But the book has two major parts. The first part explains the proposal itself (Chapters 1–6); the second part consists of an application of the proposal to Hungarian data (Chapters 7–8), and cross-linguistically (Chapter 9). Chapter 10 concludes the monograph. In what follows, I will concentrate on certain aspects of the book, such as Hunyadi's treatment of Hungarian prosody (Section **2**), his views on the connections between discourse and prosody (Section **3**), and his main proposal: the idea that stress marks scope (Section **4**). My review complements another review by Kenesei (2005) in the *Journal of Linguistics* with a somewhat different emphasis.

2. Hunyadi's theory of Hungarian prosody

Hunyadi (26–7) regards prosodic structure as not completely independent from, but also not directly dependent on, syntactic structure. (Although note that no specific syntax–prosody mapping rules are given.) His adopted framework is a metrical theory of prosody. He assumes the existence of prosodic structure made up of phonological words, prosodic phrases and intonational phrases.

As is well known, the physical characteristics of stress are rather elusive: stress often associates with pitch movement, but also with changes in the energy level. Hunyadi proposes to solve the notorious problem of stress identification and unite the effects of pitch and intensity in the following way. He defines a measure that he calls PET (pitch and energy over time), which is a number that we get at each point in the utterance, by subtracting the normalised energy value from the normalised pitch value at that point (48). As a result, the PET score at any given moment is 0 if pitch and energy change at the same rate at the point of measurement. Positive PET-values indicate that pitch is higher compared to the overall pitch of the utterance than energy is, compared to overall energy in the utterance; negative PET-values indicate the supremacy of energy over pitch. He claims that stress has the following characteristics: it starts with relative prominence of pitch, followed by an even larger relative prominence of energy, followed by a similar relative prominence of pitch (49–50). Although, at first sight, one has the impression that PET subtracts 'apples' from 'oranges', it should be noted that it does seem to provide easily readable diagrams, even for the untrained eye, which may turn out to be an advantage. But being based on discrete pitch and energy values, it will, of course, never be more than a function of these values.

In Hungarian, facts about the placement of main and secondary stresses have been a matter of some controversy (compare Kálmán–Nádasdy 1994; É. Kiss 1994; etc.) One of the most important aspects of Hunyadi's work is that he provides phonetic measurements of native speakers' recordings to support his claims. Alongside the diagrams that accompany every utterance given in the text, there is an additional database of recordings of 26 Hungarian utterances by 7 female and 7 male speakers in the Appendices. This in itself makes the book a very important source for phonetically inclined researchers of the syntax–phonology interface of Hungarian.

3. Prosody–discourse connections

It is well known that Hungarian has specialised preverbal, left-peripheral syntactic positions for topics, focus and wide-scope quantifiers. The order of elements is in (1). ("*" marks recursivity.)

(1) [Topic* [Quantifier [Focus [V...]]]

Chapter 3 addresses the question whether the pragmatic functions of focus, contrastive topic and the "neutral part" can be considered as "semantico-logical functions directly related to their prosodic realization" (55). Let us concentrate on focus, which is in any case the constituent whose analysis is the most detailed in the book.

Hunyadi aligns himself with the so-called "focus-to-accent" approach (Ladd 1996), which claims that prosodic prominence (i.e., main stress and/or corresponding accent) marks the focus of the utterance, and not the other way around. He gives two arguments for this claim. First, as (2) shows, in Hungarian, utterances containing multiple foci involve only one instance of movement, the second (and any further) focus is *in situ*. Given that both moved and *in-situ* foci bear main stress and accent, it follows that it is the prosody, rather than the position that marks focus (56–7).[1] (Capitals mark main stress.)

(2) JÁNOS olvasta a KÖNYVET.
'JOHN read the BOOK.'

The second argument concerns universal quantifiers. These are excluded from the syntactic focus position because of the incompatibility between the semantics of this position (i.e., exclusion by identification) and the universality of the quantifiers. As Hunyadi shows, notwithstanding this semantic incompatibility, the **prosodic** characteristics of a fronted (or *in-situ*) focus can also appear on fronted universal quantifiers, as in (3). So, focal meaning, in the sense of discourse pragmatics, rather than semantics, can be marked prosodically, even when syntactic marking is impossible.[2]

(3) MINDIG olvasta a könyvet.
always read the book
'He/she ALWAYS read the book.'

[1] In a later remark, Hunyadi (152) claims that the syntactic focus position indicates the presence of a "FOCUS operator", whose semantic function is "identification". A similar idea is spelt out in great detail by Horvath (2000 and subsequent work). (Hunyadi discusses the semantic characteristics of *in situ* focus in Chapter 6, Section 6.2.)

[2] A similar stress-to-focus-type proposal is put forward in Szendrői (2001); Szendrői (2003). The main difference between the two proposals is that the existence of an independent syntactic FOCUS-operator is not assumed there. Rather, it is claimed that focus movement targets the position where main stress is assigned by the rules of Hungarian prosody. In that work, no direct connections are drawn between scope assignment and prosody.

4. Prosodic scope marking

But, according to Hunyadi, it is not only the case that the discourse notion 'focus' is prosodically marked in every langauge. Rather, languages indicate operator–scope relations in general prosodically. In particular, the language universal in (4) is the central claim of this monograph (210).

(4) PRINCIPLE OF SCOPE ASSIGNMENT
Assign scope by stress.

Hunyadi takes (4) to have a direct effect on prosody. The Hungarian stress assignment rules (90) state that in the unmarked case there are "even stresses on all or almost all phonological words" (97). The prosodic operation "stress reduction" is taken to be the primary operation effecting stress assignment in Hunyadi's system.[3] The application of this operation is directly conditioned by the "operator status" of the elements involved: an operator that takes wide scope induces stress reduction on the elements in its scope (111). In other words, in Hungarian, stress on the operator and stress reduction on its scope are the consequences of semantically-driven prosodic operations.

Hunyadi (90) also assumes that there is a specific hierarchy of operators for stress reduction. This enables him to give an account of a wide set of data, encompassing interactions between the operator *csak* 'only', the FOCUS operator, sentential negation, constituent negation, emphatic *is* 'also', contrastive *is* 'even', verbal particles marking perfectivity, numerals, referentiality and modals. Even if one does not always precisely agree with the specific analyses, Chapter 8 is a great source of interesting and important observations for anyone working on the interaction of operators in Hungarian, from a semantic, syntactic or prosodic point of view.

To give just one example, Hunyadi (153–5) argues that the *csak* 'only' operator is different from the FOCUS operator and it is higher in the hierarchy of stress reduction. In an utterance like (5), this is clearly shown by the fact that the operator, which takes scope over the focal element, bears main stress, while the focus itself, which is lower in the hierarchy, undergoes stress reduction.

(5) CSAK Jánossal találkoztam.
only John-with met-I
'I only met JOHN—and noone else.' (DP focus)

However, the *csak* 'only' operator does not obligatorily attract main stress in Hungarian. The utterance in (6) is equally grammatical, in fact, on one reading, (5) and (6) are synonymous. Hunyadi argues that the reason for the optionality of stress on

[3] Another prosodic operation "neutralization" is also defined (104). This operation ensures that elements whose pragmatic status is given undergo stress reduction. The operation creates the discourse domain called "the neutral part", which follows the focus. The neutral part is prosodically unmarked, no special tune is attached to it. This puts it in contrast to focus and topic, which both have distinctive tunes. So, Hunyadi takes the position, even if he does so implicitly, that the prosodic operation associated with focus is distinct from the one associated to givenness (see also Reinhart 1995; and contra Schwarzschild 1999).

the *csak* 'only' operator is due to the fact that it is a **lexical** operator, which is therefore capable of marking its scope syntactially (in Hunyadi's (153) words, "by linear precedence"), rather than prosodically.

(6) Csak JÁNOSSAL találkoztam.
only John-with met-I
'I only met JOHN—and noone else.' (DP focus)

This, of course, means that the claim in (4) has to be weakened, as it turns out that lexical operators may mark their scope syntactically, not only prosodically (153). At the same time, we get an interesting explanation for the following fact. In contrast to the utterance in (5), which only allows the reading indicated there, the utterance in (6), repeated here in (6′), is ambiguous.[4]

(6′) Csak JÁNOSSAL találkoztam.
only John-with met-I

(a) 'I only met JOHN—and noone else.' (DP focus)
(b) 'I only met JOHN—and nothing else happened.' (IP focus)[5]

It transpires from Hunyadi's argumentation that prosodic marking on the scope of *csak* 'only' (i.e., on the focus of the utterance) is necessary to get focal ambiguity. If for some independent reason the operator is stressed, and not the focus, the ambiguity disappears. This is a very interesting point about the interaction of prosody and operators like *csak* 'only', with potential implications for issues such as the semantic debate about the focus-sensitivity of *only*.

As for the surface order of Hungarian operators (see (1) above), Hunyadi argues (116) that it follows from the proposed stress reduction rule, the left-to-right direction of its application, and the assumed hierarchy of the operators for stress reduction. This proposal represents an original approach to the Hungarian left periphery. In fact, it is probably a unique attempt to explain the order of elements, rather than just stating it. Nevertheless, I think certain issues remain unsolved. What I believe to be the most important one concerns the assumed hierarchy of operators. It seems to me that many of the derived orderings rely heavily on this hierarchy. At the same time, the hierarchy itself is to a large extent not independently motivated, although Hunyadi shows that at least certain aspects of it are valid cross-linguistically (see Chapter 9).

A more concrete issue concerns topics. Non-contrastive, unstressed topics such as *a postást* 'the postman-acc' in (7) indeed receive wide scope as their syntactic position suggests, but they do not induce destressing of any constituents in their scope, contrary to what one would expect under the proposed theory. On the other hand, contrastive topics, as in (8), receive their own main prominence within their own intonational

[4] Note that although Hunyadi claims that subtle prosodic differences may distinguish the two readings, I do not think this would be true in general.

[5] Hunyadi calls this reading "VP-focus", I prefer the indicated "IP focus" or "all-focus". But this is not important, as his point is simply that wide focus is possible here; whether it is VP or IP is irrelevant.

phrase. Contrastive topics, arguably, take narrow scope with respect to a following focus. According to Hunyadi (63–5), the reason why it is the focus rather than the topic that takes wider scope is that the intonational tune attached to the focus is an independent one, while the tune on the topic suggests incompleteness. This may be so, in fact this is probably very close to the truth, but nevertheless, this is an extra assumption that the author has to take on board to derive the order in (1).

(7) A postást megharapta a kutya.
the postman-acc prt-bit the dog
'The dog bit the postman.'

(8) A postást, megharapta a kutya. A szomszédot, nem.
the postman-acc prt-bit the dog the neighbour-acc not
'The postman was bitten by the dog. The neighbour wasn't.'

As far as the cross-linguistic generalisation of the theory is concerned, in Chapter 9, Hunyadi states that there is parametric variation amongst languages as to the way the principle in (4) gets realised:

(9) (a) Assign scope by stressing the scope.
(b) Assign scope by reducing the stress of the scope.

Let us concentrate here on the discussion on Finnish. Hunyadi (212) claims that Finnish subscribes to 'stressing the scope' while, we may recall that Hungarian 'reduces stress of the scope'. In addition, he provides data showing that Finnish does not have fronted unstressed topics and that it is not possible to distinguish prosodically a contrastive topic from a focus in this language. He claims that these facts follow from the idea that Finnish marks scope by stress, rather than by reducing stress, as Hungarian does.

In my opinion, the data that he gives to support this position is not sufficiently detailed. It is, of course, completely understandable that the empirical scope of the proposal is more limited in the case of the languages considered here than in the case of Hungarian. However, this makes it difficult to evaluate the proposal. For instance, there seems to be "a phonological rather than logical requirement for a sentence-initial word to bear stress" in Finnish (215). It is not clear to me how this claim can be incorporated into the system of stress assignment that Hunyadi proposes. It is also not clear whether this claim is compatible with another claim made elsewhere, stating that in Finnish, the direction of stressing is to the right (211). This is important, because the presence of an obligatory main stress on the leftmost phonological word may actually provide an alternative explanation to the unavailability of unstressed topic fronting in the language.

5. Conclusion

I would like to conclude that the innovativity of Hunyadi's approach to the interactions between semantics and prosody and its extensive phonetic experimentation is very much to be appreciated, even if I sometimes disagree with the details of his analysis on specific data.

On a more text-oriented note, given the complex nature of the subject matter of the book, it would have benefited from more rigorous editing. (Although it must be noted that the abstracts at the beginning of each chapter are of great help to the reader.) There are missing glosses (e.g., for the Finnish examples (289–91)) and sometimes one has the impression that the phonetic diagrams may have been accidentally mismatched with their descriptions in the text (e.g., (18a) is clearly not the same as (18); example (36) and the text immediately after it).

Kriszta Szendrői

References

Büring, Daniel 1997. The meaning of topic and focus—The 59th Street Bridge accent. Routledge, London.

É. Kiss, Katalin 1994. Sentence structure and word order. In: Katalin É. Kiss – Ferenc Kiefer (eds): The syntactic structure of Hungarian (Syntax and semantics 27), 1–90. Academic Press, New York.

Frascarelli, Mara 2000. The syntax–phonology interface in focus and topic constructions in Italian. Kluwer, Dordrecht.

Horvath, Julia 2000. Interfaces vs. the computational system in the syntax of focus. In: Hans Bennis – Martin Everaert (eds): Interface strategies, 183–206. Holland Academic Graphics, Amsterdam.

Kálmán, László – Ádám Nádasdy 1994. A hangsúly [Stress]. In: Ferenc Kiefer (ed.): Strukturális magyar nyelvtan 2: Fonológia [A structural grammar of Hungarian 2: Phonology], 393–467. Akadémiai Kiadó, Budapest.

Kenesei, István 2005. Hungarian in focus (Review article). In: Journal of Linguistics 41 : 409–35.

Ladd, Robert 1996. Intonational phonology. Cambridge University Press, Cambridge.

Reinhart, Tanya 1995. Interface strategies (OTS Working Papers in Theoretical Linguistics). OTS, Utrecht.

Schwarzschild, Roger 1999. GIVENness, Avoid F and other constraints on the placement of focus. In: Natural Language Semantics 7 : 141–77.

Szendrői, Kriszta 2001. Focus and the syntax-phonology interface. Doctoral dissertation, University College London. (http://www.phon.ucl.ac.uk/home/kriszta/)

Szendrői, Kriszta 2003. A stress-based approach to the syntax of Hungarian focus. In: The Linguistic Review 20 : 37–78.

Zubizarreta, Maria-Luisa 1998. Prosody, focus, and word order. MIT Press, Cambridge MA.

Acta Linguistica Hungarica, Vol. 53 (3), pp. 341–342 (2006)

HUNGARIAN BOOKS ON LINGUISTICS

Ferenc Kiefer Péter Siptár (eds): Magyar nyelv [The Hungarian language]. Akadémiai Kézikönyvek, Akadémiai Kiadó, Budapest, 2006. 1111 pp.

This is a new, expanded, revised and updated version of *A magyar nyelv kézikönyve* (cf. vol. 51, p. 223 of the present journal). The new chapters cover topics like semantics, pragmatics, a typological overview of the Uralic languages, the Modern Hungarian period, the history of linguistics in Hungary (prior to the 20th century), speech defects, the Hungarian Sign Language, language and writing, foreign language teaching, etc. Some of the "old" chapters have also been substantially re-written and expanded.

Contents: Introduction (Ferenc Kiefer); Part one: Linguistics and its areas (Ádám Nádasdy), Phonology (Péter Siptár), Morphology (Ferenc Kiefer), Parts of speech (István Kenesei), Syntax (Katalin É. Kiss), Text linguistics (Gábor Tolcsvai Nagy), Semantics (Ferenc Kiefer – Beáta Gyuris), Pragmatics (Enikő Németh T.);

Part two: Typological characteristics of the Uralic languages (Marianne Bakró-Nagy), Hungarian as a Uralic language (Katalin Sipőcz), The Ancient Hungarian period (László Horváth), The Old Hungarian period (Klára Korompay), The Middle Hungarian period (Lea Haader), The Language Reform (Adrienne Dömötör), The Modern Hungarian and Contemporary Hungarian periods (Krisztina Laczkó), Word stock (Károly Gerstner), The history of linguistics in Hungary from the beginnings to the early twentieth century (András Cser);

Part three: Dialects and regional varieties (Jenő Kiss), Varieties of Hungarian spoken in other countries (Miklós Kontra), On the variability of language use within Hungary (Miklós Kontra), Bilinguality and multilinguality (Anna Borbély), Stylistics (Gábor Tolcsvai Nagy), Neurolinguistics (Zoltán Bánréti), Psycholinguistics (Csaba Pléh), Child language (Csaba Pléh);

Part four: Phonetics (Ilona Kassai), Speech defects (Mária Gósy), The Deaf community in Hungary and the Hungarian Sign Language (Csilla Bartha Helga Hattyár – Mária Helga Szabó), Language and writing (Ádám Nádasdy), Hungarian for special

1216–8076/$ 20.00

purposes (Zsuzsa Kurtán), Language planning, language policy, language cultivation (Klára Sándor), Foreign language learning, foreign language teaching (Dorottya Holló), Language and the law (Miklós Kontra), Language technologies (Gábor Prószéky – Gábor Olaszy – Tamás Váradi); Abbreviations, Names Index, Subject Index.

INFORMATION FOR CONTRIBUTORS

Please provide typewritten double-spaced manuscripts with wide margins (50 characters per line, 25 lines per page). Type on one side of the paper only and number all the pages consecutively. The following information should be provided on the first page: the title of the article (or, in case of reviews, the title of the book under discussion, the name of the publisher, the place and date of publication as well as the number of pages), the author's full name and address (including email address) and an abbreviated title of the paper (not exceeding 45 characters including spaces). Only original papers will be published and a copy of the *Publishing Agreement* will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement. Authors are requested to send two hard copies of their manuscript + a floppy disk under DOS/Windows. The manuscript should be accompanied by a list of five keywords and an abstract of about 100 words. It should be typed on a separate sheet of paper. Tables, diagrams and illustrations (with the author's name and an Arabic number) should be presented on separate sheets. Captions should be typed on a separate sheet and placed at the end of the manuscript. Footnotes should be typed double-spaced on a separate sheet at the end of the article. All language examples (and only these) are to be italicized (single underlining). Citations in the text should be enclosed in double quotation marks (" " in case of a paper written in English, „ " in German and « » in French).

Reference to a publication should be made by the name of the author, the year of publication and, when necessary, by page numbers in the following ways:

...as described by Schmidt (1967)...
...as referred to by Hoover (1967, 56–78; 1976, 43)...
...mentioned by several authors (Elgar 1978, 77; Williams 1981, 154–6).

An alphabetically arranged list of references should be presented at the end of the article as follows:

Bárczi, Géza 1958a. Magyar hangtörténet [The historical phonology of Hungarian]. Akadémiai Kiadó, Budapest.

Bárczi, Géza 1958b. A szótövek [Word stems]. Akadémiai Kiadó, Budapest.

Brockhaus, Wiebke 1995. Skeletal and suprasegmental structure within Government Phonology. In: Jaques Durand – Francis Katamba (eds): Frontiers in phonology: Atoms, structures, derivations. 180–221. Longman, Harlow.

Cole, Jennifer 1995. The cycle in phonology. In: Goldsmith (1995: 206–44).

Goldsmith, John A. (ed.) 1995. The handbook of phonological theory. Blackwell, Cambridge MA & Oxford.

Kaye, Jonathan – Jean Lowenstamm – Jean-Roger Vergnaud 1990. Constituent structure and government in phonology. In: Phonology 7: 301–30.

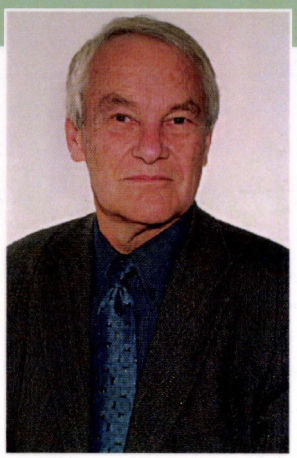
Tomioka, Satoshi 1997. Focusing effects and NP-interpretation in VP-ellipsis. Doctoral dissertation, University of Massachusetts, Amherst.

For marking subsections decimal notation should be applied. Do not use more than four digits if possible. Examples within the text should be marked in italics. Meanings are to be rendered between inverted commas (' '). If glosses are given morpheme by morpheme, the initial letter of the gloss should be placed exactly below that of the example. Grammatical morphemes can be abbreviated in small case letters connected to the stem or the other morphemes by a hyphen. No period should be applied in such abbreviation. For example:

(1) (a) A sólymaid elszálltak
the falcon-gen-pl-2sg away-flew-3pl
'Your falcons have flown away.'

Examples can be referred to in the text as (1a), (1a–d), etc.

One proof will be sent to the author. Please read it carefully and return it by air mail to the editor within one week.



FERENC KIEFER, editor-in-chief, professor emeritus
Member of the Hungarian Academy of Sciences
President of CIPL
Corresponding member of the Austrian Academy of Sciences
Member of Academia Europaea and of
European Academy of Arts, Sciences and Humanities
Main research areas:
morphology, lexical semantics, semantics-pragmatics interface
He has published 11 books (in Hungarian, English, French and German) and more than 200 articles

CODEN: ACTA LINGUIST. HUNG.
SIGILLUM: ALHUE 8

ISSN 1216-8076

Volume 53 ▪ Number 4 ▪ December 2006

Editor-in-Chief ▪ FERENC KIEFER

FOUNDED IN 1951

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

WWW.AKADEMIAI.COM

AKADÉMIAI KIADÓ

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Acta Linguistica Hungarica publishes papers on general linguistics with particular emphasis on discussions of theoretical issues concerning Hungarian and other Finno-Ugric languages. Papers presenting empirical material must have strong theoretical implications.
The scope of the journal is not restricted to the traditional research areas of linguistics (phonology, syntax and semantics, both synchronic and diachronic), it also covers other areas such as socio- and psycholinguistics, neurolinguistics, discourse analysis, the philosophy of language, language typology, and formal semantics.
The articles should be written in English, exceptionally, however, papers written in German and French are also accepted. The journal is published in yearly volumes of four issues.

■

Abstracted/indexed in

Bibliographie Linguistique/Linguistic Bibliography, International Bibliographies IBZ and IBR, Linguistics Abstracts, Linguistics and Language Behaviour Abstracts, MLA International Bibliography.

■

Editorial correspondence

ACTA LINGUISTICA HUNGARICA
Research Institute for Linguistics
P.O. Box 701/518, H-1399 Budapest, Hungary
Phone: (+36 1) 351 0413; Fax: (+36 1) 322 9297
E-mail: kiefer@nytud.hu; siptar@nytud.hu

■

Subscription price

for Volume 53 (2006) in 4 issues EUR 252 + VAT (for North America: USD 316) including online access and normal postage; airmail delivery EUR 20 (USD 25).

■

Please send your order to

AKADÉMIAI KIADÓ
Journals Business Centre
P.O. Box 245, H-1519 Budapest, Hungary
Phone: (+36 1) 464 8222; Fax: (+36 1) 464 8221
E-mail: journals@akkrt.hu
www.akademiai.com; www.akademiaikiado.hu

■

ISSN 1216-8076

ALing 53 (2006) 4

Printed in Hungary

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Volume 53, Number 4, December 2006

AKADÉMIAI KIADÓ
MEMBER OF WOLTERS KLUWER GROUP

Cover design: xfer grafikai műhely

CONTENTS

Acta Linguistica Hungarica, Vol. 53 (4), pp. 343–391 (2006)
DOI: 10.1556/ALing.53.2006.4.1

WORD ORDER IN HUNGARIAN EXCLAMATIVES*

ANIKÓ LIPTÁK

Centre for Linguistics
Leiden University
P. O. Box 9515
2300 RA Leiden
The Netherlands
A.Liptak@let.leidenuniv.nl

Abstract: This paper gives a syntactic overview and analysis of exclamative constructions in Hungarian. Its main purpose is to describe word order variation in exclamative clauses, in comparison with other sentence types. The formal properties of exclamatives that will be discussed here have important consequences for the theories of exclamatives and exclamativity in general. The empirical findings will force one to reconsider the syntactic theory of exclamatives put forward by Portner and Zanuttini (2003). The key modification affects the role focus plays in exclamatives: it will be shown that languages can use available syntactic means of focusing in the expression of exclamatives.

Keywords: exclamatives, focus, *wh*-pharse, variable, speech act

1. Introduction to exclamatives

Exclamative sentences are a sentence type used to express surprise or astonishment about something that is unexpected or extraordinary. Unlike declaratives and interrogatives, exclamatives are considered a **minor**

* I hereby thank the audiences of the 30th Incontro di Grammatica Generativa and the Seventh International Congress of Hungarian Linguistics, as well as Marcel den Dikken, István Kenesei and an anonymous reviewer for helpful comments on the material presented in this paper and in Lipták (2005). All remaining errors are of course my own. This research was supported by NWO (Netherlands Organization for Scientific Research).

1216–8076/$ 20.00

sentence type (Sadock–Zwicky 1985), because not all languages possess exclamative sentences that are formally distinguishable from other sentence types.

Exclamations in Hungarian can be expressed in various ways. The most standard of these are construction types in which there is an exclamative phrase that expresses the surprising thing or property.[1] This exclamative phrase (referred to as the "E-phrase" in the following) can be (i) a wh-phrase, (ii) a phrase with *de*, and (iii) a phrase formed by relativization. According to these, we can speak about *wh*-exclamatives, *de*-exclamatives and relativized exclamatives respectively. The following examples illustrate the three distinct types:[2]

(1) [E Mennyi könyvet] elolvastál! *wh*-exclamative
how.many book-acc pv-read-2sg
'You read so many books!'

(2) [E De sok könyvet] elolvastál! *de*-exclamative
de many book-acc pv-read-2sg
'You read so many books!'

(3) [E Amennyi könyvet te elolvastál]! relativized exclamative
rel-how.many books-acc you pv-read-2sg
'The number of books you read!'

[1] In addition to this type there also exist exclamatives which express surprise about the polarity of a proposition, like *Is she cute!* in English. Such propositional exclamatives will not be discussed here, as Hungarian does not code these in a syntactically interesting way.

[2] The following glosses are used in this article: ' = emphatic stress; acc = accusative case; Adj = adjective; Adv = adverb; dat = dative case; E = exclamative; N = noun; pl = plural; pv = preverb(al element); rel = relative morpheme. Nominative case is not glossed. Subject person and number morphemes are 1/2/3sg/pl; tense and definiteness agreement morphemes are not glossed. SMALL CAPS on lexical words indicate contrastive focus. Hungarian examples are translated into English using English exclamatives when possible. When this is not possible (English has a fewer range of acceptable E-phrases), the English translation will be given with a *so...* or *such a...* phrase or will be embedded under an exclamative predicate.

As their translations indicate, these three types do not differ in meaning: they all indicate the surprise of the speaker about something outstanding: the unexpectedly high number of books that the addressee read.[3]

The main focus of this study is on the word order properties of the above types of exclamatives, concerning the syntactic distribution of exclamative phrases in them, also in comparison with other sentence types (indicatives and interrogatives). It will be shown that the placement of exclamative phrases follows well-defined rules that are distinct from that of interrogative phrases. Interpretation-wise, all exclamative phrases are focused with a scalar focus reading, and accordingly, their placement is into one of the positions where focal constituents can appear. Which focal position is selected is determined by the lexical properties of the E-phrase itself. These findings provide new insights for theories of the exclamative sentence type cross-linguistically.

The structure of the paper is as follows. Section **2** presents the data of Hungarian exclamatives that will be analyzed in later sections. The discussion is confined to *wh*-exclamatives and *de*-exclamatives only. Section **3** spells out the syntactic properties of the observed data, sketching the syntactic position for all types of E-phrases in the language. Section **4** provides a semantic and syntactic analysis of the observed patterns, touching also on the syntax of the relativized exclamative type in example (3) above. Section **5** summarizes the findings of the paper and spells out the theoretical consequences of these for the theory of exclamatives recently put forward in Portner–Zanuttini (2003). It will be shown that Portner and Zanuttini's theory needs to be amended to allow for non-*wh*-exclamative phrases and unembeddable exclamatives.

2. Syntactic properties of Hungarian exclamatives

2.1. Some basic properties of *wh*- and *de*-exclamatives

Wh- and *de*-exclamatives contain an exclamative constituent formed with a *wh*-phrase or a *de*-phrase. Before turning to the distribution of these

[3] The three exclamative sentence types differ in their prosody. (1) and (2) have stress on the E-phrase and falling intonation following it (which might be sharper in the case of (2)). Sentence (3) has main stress on the verb and falling intonation characterizes only the very end of the sentence.

phrases, some discussion is in order about the general properties of *wh*- and *de*-exclamatives.

According to my small survey among 10 Hungarian speakers, *wh*-exclamatives can make use of any *wh*-word that can occur in Hungarian questions. Next to (1) above, an illustrative bunch of other examples is given in (4) below. As can be seen from the translations, the *wh*-phrases in such exclamatives do not refer to an individual variable that the speaker cannot identify, as in the case of questions, but to a degree expression, which is associated with a high scalar value (Elliott 1974). Note also that *wh*-exclamatives, even when they occur as root clauses can always be introduced by the regular finite complementizer *hogy* 'that'. The presence of such a complementizer adds extra (emotional) emphasis to the exclamative utterance as a whole.

(4) (a) (Hogy) ki jött el ebbe a faluba!
comp who came-3sg pv this-into the village-into
'What a person came to this village!' (scale: properties of people)

(b) (Hogy) mi esett meg ebben a faluban!
comp what happened-3sg pv this-in the village-in
'What a thing happened in this village!' (scale: properties of events)

(c) (Hogy) hova bújtak a gyerekek!
comp where hid-3pl the children
'In what strange places the children hid!' (scale: properties of places)

(d) (Hogy) mikor jöttél tegnap haza!
comp when came-2sg yesterday home
'At what strange time you came home yesterday!' (scale: properties of times)

(e) (Hogy) melyik könyvet vetted meg!
comp which book-acc bought-2sg pv
lit. '(I am surprised at) which book you bought!' (scale: properties of books)

(f) (Hogy) milyen ruhában mentél dolgozni!
comp what.kind cloth-IN went-2sg work-infe
'The kind of clothes you went to work in! (scale: properties of clothes)

(g) (Hogy) hogy egyensúlyozott Béla a biciklin!
comp how balanced-3sg Béla the bike-on
'How Béla was balancing on the bike!
(scale: properties of manners of balancing)

As Kálmán (2001) mentions,[4] *miért* 'why' is exceptional in that it cannot occur in exclamatives. According to my findings, this is subject to individual variation. Some speakers accept *miért* in exclamatives, others do not. Note that corresponding nominal phrases like *milyen furcsa okból* 'for what a strange reason' is perfectly fine for all speakers (5b):

(5) (a) %(Hogy) te miért hívtad fel Annát!
comp you why called-2sg pv Anna-acc

(b) (Hogy) te milyen furcsa okból hívtad fel Annát!
comp you what.kind strange reason-from called-2sg pv Anna-acc
'For what a strange reason you called Anna!'

When it comes to possible and impossible *wh*-E-phrases, Kálmán (*ibid.*) notes that *wh*-phrases formed with *is* 'also', as well as aggressively non-D-linked expressions cannot be used in exclamatives (6a), (7a). These are of course perfectly fine in questions (6b), (7b):

(6) (a) *Ki is ment el! exclamative
who also went pv

(b) Ki is ment el? question
who also went pv
'Who was it again who left?'

(7) (a) *Ki a fene ment el! exclamative
who the hell went pv

(b) Ki a fene ment el? question
who the hell went pv
'Who the hell left?'

The ungrammaticality of the exclamative examples (6a), (7a) derives from the fact that exclamatives are **factive** (Grimshaw 1979, see also section 4 below): their propositional content is presupposed. The content of (6b) and (7b) cannot be presupposed, as is shown by the fact that the questions in (6b), (7b) cannot be embedded under factive predicates.

(8) (a) *Tudom, hogy ki is ment el.
know-sg that who also went pv
'I know who it was again who left.'

[4] The section on exclamatives in Kálmán (2001) was authored by Viktor Trón.

(b) *Tudom, hogy ki a fene ment el.
know-sg that who the hell went pv
'I know who the hell left.'

It is therefore not surprising that such *wh*-expressions cannot occur in exclamatives, either.

Just as there exist *wh*-phrases that cannot occur in exclamatives, there are also *wh*-phrases which can only occur in exclamatives. *Wh*-phrases involving strong evaluative adjectives or adverbs like *rohadtul* 'rottenly' for example can occur in E-phrases (Kálmán *ibid.*):[5]

(9) (a) Milyen rohadtul megfáztam!
how rottenly pv-cold.caught-1sg
'What an awful cold I got!'

(b) *Milyen rohadtul fáztál meg?
how rottenly cold.caught-2sg pv
'How very badly did you catch a cold?'

This is due to the presence of the evaluative adverb, which is used to express the speaker's strong judgement about the cold he got. In the case of (9b), where the *wh*-constituent denotes a variable unknown to the speaker, the same evaluative judgement cannot be cast.

Another property of *wh*-exclamatives is that they can be embedded under exclamative predicates:

[5] Some *wh*-items modified by *minden* 'all' and *-csoda* 'wonder' affixed *wh*-items are also typical of exclamatives only:

(i) (a) Hol mindenhol kiöntött a Tisza!
where everywhere pv-flooded the Tisza
'The Tisza flooded at so many places!'

(b) ???*Hol mindenhol öntött ki a Tisza?
where everywhere flooded pv the Tisza
'Which were all the places where the Tisza flooded?'

(ii) (a) Micsoda képek vannak a múzeumban!
what-wonder pictures are the museum-in
'What beautiful pictures there are in the museum!'

(b) %Micsoda képek vannak a múzeumban?
what-wonder pictures are the museum-in
'What (kind of) pictures are there in the museum?'

(10) (a) Elképesztő, hogy milyen rohadt hideg van.
astonishing that how rotten cold is
'It's astonishing how awfully cold it is.'

(b) Meglep, hogy milyen rohadt hideg van.
surprise-3sg that how rotten cold is
'It surprises me how awfully cold it is.'

Turning now to the other type of exclamatives, *de*-exclamatives, the first thing to be mentioned is that they, too, are associated with a high scalar value, just like *wh*-exclamatives. *De*-exclamatives are formed with the word *de*, which is homophonous with the adversative coordinator *de* 'but'. Categorically, *de* distributes as *milyen* 'what kind', modifying an adjective or an adverb (11a, b, c). In another distribution *de* is a VP-adverb, indicating degree or intensity of the event (11d):

(11) (a) {De / Milyen} piszkos ruhában mentél dolgozni!
de how filthy cloth-in went-2sg work-inf
'You went to work in such filthy clothes!'

(b) {De / Milyen} ügyesen egyensúlyozott Béla a biciklin!
de how skillfully balanced-3sg Béla the bike-on
'How skilfully Béla was balancing on the bike!'

(c) {De / Milyen} sok könyvet elolvastál!
de how many book-acc pv-read-2sg
'You read so many books!'

(d) De becsaptad az ajtót!
de pv-slammed-2sg the door-acc
'How strongly you slammed the door!'

De-phrases cannot contain a *wh*-word. In other words, we cannot find an exclamative that is both a *de*-exclamative and a *wh*-exclamative at the same time:

(12) *De piszkos miben mentél dolgozni!
de filthy what-in went-2sg work-inf
'The filthy things you went to work in!'

De-exclamatives sharply differ from *wh*-exclamatives when it comes to compatibility with an overt complementizer and an embedding predicate. Unlike *wh*-exclamatives, *de*-exclamatives cannot be introduced by a finite complementizer in root contexts and cannot be embedded under a matrix

exclamative predicate. Compare the following examples with examples (4) and (10) above:

(13) (a) (*Hogy) de rohadt hideg van!
comp *de* rotten cold is
'How awfully cold it is!'

(b) *Elképesztő, hogy de rohadt hideg van.
astonishing that *de* rotten cold is
'It's astonishing how awfully cold it is.'

(c) *Meglep, hogy de rohadt hideg van.
surprise-3sg that *de* rotten cold is
'It surprises me how awfully cold it is.'

These facts will be further commented on in section 4.

2.2. The distribution of exclamative phrases in *wh-* and *de*-exclamatives

As the careful reader presumably has noticed already, in all examples with *wh-* or *de*-exclamatives above, the E-phrase appears in a preverbal position. Postverbal occurrences of exclamative phrases are ungrammatical:

(14) *Elolvastál mennyi könyvet / de sok könyvet!
pv-read-2sg how.many book-acc *de* many book-acc
'You read so many books!'

The preverbal position in which E-phrases can be found is furthermore not just any preverbal position. E-phrases in this position must always be adjacent to the verb and are obligatorily marked with heavy stress:

(15) (a) (Hogy) melyik könyvet (*tegnap) vetted meg!
comp which book-acc yesterday bought-2sg pv
lit. '(I am surprised at) which of the books you bought yesterday!'

(b) De sok könyvet (*tavaly) elolvastál!
de many book-acc last year pv-read-2sg
'You read so many books last year!'

The adjacent preverbal position that E-phrases occupy can be of two types, depending on the presence or absence of inversion between the verb and the preverbal particle, if the verb has the latter. In one pattern,

the E-phrase is accompanied by preverb-verb inversion (16a), in the other pattern preverb and verb show the straight (uninverted) order (16b):

(16) (a) (Hogy) mi esett meg ebben a faluban!
comp what happened-3sg pv this-in the village-in
'What a thing happened in this village!'

(b) (Hogy) mennyi könyvet elolvastál!
comp how.many book-acc pv-read-2sg
'You read so many books!'

Presence or absence of inversion is not an arbitrary property of exclamative sentences. As section **3** below will show, the two word order patterns correspond to a clear difference in meaning, and, accordingly, they are selective of the kind of E-phrases they can occur with. In my findings, Hungarian differentiates between three types of E-phrases according to their behaviour in exclamatives. Some E-phrases can:

(i) only occur with the straight order (without inversion)
(ii) occur both with and without inversion
(iii) only occur with inversion

The following subsections give an illustrative characterization of each type in turn. It is important to stress that the classification to be provided is by no means exhaustive, when it comes to each and every possible E-phrase. The goal is rather to show the basic patterns, which will serve as the basis of the theoretical discussion in section **3**.

2.2.1. E-phrases that only occur with the straight order

The group of E-phrases that under all circumstances have to occur with straight pv-V order involve the *wh*-phrases *mennyire* 'to what extent/how much', and *hogy* lit. 'how', in the meaning 'to what extent/how much':

(17) (a) (Hogy) mennyire {megnőtt / *nőtt meg} Éva!
comp how.much pv-grew-3sg grew-3sg pv Éva
'How much Éva has grown!'

(b) (Hogy) hogy {megnőtt / *nőtt meg} Éva!
comp how.much pv-grew-3sg grew-3sg pv Éva
'How much Éva has grown!'

Apart from these *wh*-expressions, *de/milyen*-phrases formed with **grade, completion** or **intensity adverbs** (adverbs expressing high or maximal

degree) show the same behaviour. Grade, completion and intensity adverbials (group C adverbs in Kiefer 1967) are for example *nagyon* 'very', *egészen* 'entirely', *alaposan* 'thoroughly', *túlzottan* 'excessively', *gyökeresen* 'radically', *mérhetetlenül* 'immensely', *kereken* 'in plain terms', *rémesen* 'dreadfully'. Combined with *milyen* or *de* they form E-phrases which can only occur without inversion:

(18) De/milyen nagyon {megnőtt / *nőtt meg} Éva!
de how much pv-grew-3sg grew-3sg pv Éva
'How much Éva has grown!'

(19) De/milyen alaposan {megfázott / *fázott meg} Ágnes!
de how thoroughly pv-cold.caught-3sg cold.caught-3sg pv Ágnes
'What a thorough cold Ágnes got!'

Grade, completion and intensity adverbs show the same syntactic behaviour in neutral indicative clauses as well: they do not trigger inversion. At the same time, they are always adjacent to the verb and they receive the main stress of the sentence:[6]

(20) (a) Éva nagyon {megnőtt / *nőtt meg}.
Éva much pv-grew-3sg grew-3sg pv
'Éva has grown a lot.'

(b) Ágnes alaposan {megfázott / *fázott meg}.
Ágnes thoroughly pv-cold.caught-3sg cold.caught-3sg pv
'Ágnes caught a thorough cold.'

2.2.2. E-phrases with optional inversion

E-phrases with optional inversion constitute a varied class, some of whose members are quite marked. Prototypical E-phrases with optional inversion are *wh*-phrases formed with *hány* 'how many', and *mennyi* 'how much/many' (and their derivatives, like *hányszor/mennyiszer* 'how often', but not *hanyadik*, the ordinal form of 'how many'). In a similar

[6] This is not the case in non-neutral indicative clauses. When the sentence contains a lexical focus, the adverbial phrase can be postverbal:

(i) ÁGNES fázott meg alaposan.
Ágnes cold.caught-3sg pv thoroughly
'It was Ágnes who caught a thorough cold.'

fashion, E-phrases containing the quantifier *sok* 'much/many' show optional inversion (21b), similarly to its behaviour in declarative clauses (21c):

(21) (a) (Hogy) hány könyvet {megvettél / vettél meg}!
comp how.many book-acc pv-bought-2sg bought-2sg pv
'You bought so many books!'

(b) {De /milyen} sok könyvet {megvettél / vettél meg}!
de how many book-acc pv-bought-2sg bought-2sg pv
'You bought so many books!'

(c) Sok könyvet {megvettél / vettél meg}.
many book-acc pv-bought-2sg bought-2sg pv
'You bought many books.'

A somewhat more marked case involves phrases with the universal quantifier *minden(ki)*. These are judged less than perfect when they occur with inversion:

(22) (a) (Hogy) ki mindenki {eljött / ?jött el} az ünnepségre!
comp who everyone pv-came-3sg came-3sg pv the celebration-to
'The (different) kinds of people/the number of people who came to the celebration!'

(b) (Hogy) mi mindent {megettél / ?ettél meg}!
comp what everything-acc pv-ate-2sg ate-2sg pv
'The number of things you have eaten!'

Finally, plural noun phrases like *kik* 'who-pl', *mik* 'what-pl', as well as plural marked *milyen (N)* 'what.kind N' phrases also occur both with and without inversion. In the case of these elements, the straight order is more marked. Half of my speakers judged these ungrammatical.

(23) (a) (Hogy) kik {%eljöttek / jöttek el} az ünnepségre!
comp who-pl pv-came-3pl came-3pl pv the celebration-to
'The kind of people who came to the celebration!'

(b) (Hogy) miket {%megettél / ettél meg}!
comp what-pl-acc pv-ate-2sg ate-2sg pv
'The things you have eaten!'

(c) (Hogy) milyen nagy házakat {%megvettetek / vettetek meg}!
comp what.kind big house-pl-acc pv-bought-2pl bought-2pl pv
'You have bought such big houses!'

Some adverbial phrases expressing manner, time and frequency (group A in Kiefer 1967) also occur with optional inversion. These are adverbs like *durván* 'in a rough manner', *gyorsan* 'quickly', *szépen* 'nicely', *bátran* 'bravely', *boldogan* 'happily', *okosan* 'wisely', *sűrűn* 'often', *korán* 'early', *hamar* 'soon', *gyakran* 'often'.

(24) {De/ milyen} durván {odaszólt / szólt oda} Ákos Ildikónak!
de how roughly pv-called-3sg called-2sg pv Ákos Ildikó-dat
'Ákos addressed Ildikó in such a rough manner!'

Such adverbials can also occur with both orders in neutral indicative contexts. With both word orders they are adjacent to the verb and can carry stress:

(25) 'Durván {odaszólt / szólt oda} Ákos Ildikónak.
roughly pv-called-3sg called-3sg pv Ákos Ildikó-dat
'Ákos addressed Ildikó in a rough manner.'

2.2.3. E-phrases with obligatory inversion

E-phrases that occur with inversion comprise all *wh-* and *de*-phrases that were not listed above in the other two types. The following *wh-* and *de*-phrases and their derivatives belong here: singular *ki(csoda)* 'who' and *mi(csoda)* 'what', *mikor* 'when', *hol/merre* 'where', *hogy(an)* 'how', *miért* 'why', *melyik (N)* 'which (N)', *milyen (N)* 'what kind of (N)', *milyen (Adj N)/ de (Adj N)* 'how (Adj N)':

(26) (a) (Hogy) ki {ment el / *elment} moziba Annával!
comp who went-3sg pv pv-went-3sg cinema-to Anna-with
'The person who went to the cinema with Anna!'

(b) (Hogy) hova {mentél el / *elmentél}!
comp where went-2sg pv pv-went-2sg
'The place you went to!'

(c) (Hogy) milyen drága könyvet {vettél meg / *megvettél}!
comp how expensive book-acc bought-2sg pv pv-bought-2sg
'How expensive a book you bought'

(d) De drága könyvet {vettél meg / *megvettél}!
de expensive book-acc bought-2sg pv pv-bought-2sg
'How expensive a book you bought!'

E-phrases containing the quantifier *kevés* 'little/few' also show obligatory inversion (27a), just like phrases with *kevés* in neutral indicative sentences (27b):[7]

(27) (a) {De / milyen} kevés könyvet {olvastál el / *elolvastál}!
de how few book-acc read-2sg pv pv-read-2sg
'How few books you read!'

(b) Kevés könyvet {olvastál el / *elolvastál}.
few book-acc read-2sg pv pv-read-2sg
'You read few books.'

Among adverbial phrases, there are two types of lexical items that obligatorily occur with inversion in exclamatives. One are adverbials headed by **exclusive** adverbs (group B in Kiefer 1967), for example, *későn* 'late', *bonyolultan* 'in a complicated manner', *hibásan* 'with mistake', *hasztalan* 'in vain', *rendetlenül* 'in a disorderly way', *nehezen* 'with difficulty', *izléstelenül* 'tastelessly'. As (28b) indicates, such adverbial phrases occur with the same word order in indicative sentences as well. They carry stress and occur adjecent to the verbal head in neutral indicatives:

(28) (a) {De / milyen} későn {feküdtél le / *lefeküdtél}!
de how late went.to.bed-2sg pv pv-went.to.bed-2sg
'How late you went to bed!'

(b) [']Későn {feküdtél le / *lefeküdtél}.
late went.to.bed-2sg pv pv-went.to.bed-2sg
'You went to bed late.'

The other class of adverbials that force inversion in exclamatives are some manner, temporal and frequency adverbs, like *lassan* 'slowly' or *barátságosan* 'in a friendly way':

(29) {De / milyen} lassan {érett meg / *megérett} a szilva!
de how slowly ripened-3sg pv pv-ripened-3sg the plum
'The plums got ripe so slowly!'

[7] Obligatory inversion can only be found in sentences without focus. When the sentence contains a lexical focus, the *kevés* phrase can be postverbal, as shown in (i). Compare also footnote 6 for similar facts.

(i) PÉTER olvasott el kevés könyvet.
Péter read-3sg pv few book-acc
'It was Péter who read few books.'

These types of adverbs have the characteristic property that although they can occur without inversion in neutral indicative sentences, in such occurrences they cannot be stressed and they often have a different meaning (30a). When these adverbs are stressed, inversion is the only order they can occur with:

(30) (a) Lassan megérett a szilva.
slowly pv-ripened-3sg the plum
'Slowly, the plums got ripe. (It is possible that the ripening itself went quickly.)'

(b) 'Lassan {érett meg / *megérett} a szilva.
slowly ripened-3sg pv pv-ripened-3sg the plum
'The plums underwent slow ripening. (It is not possible that the ripening itself went quickly.)'

2.3. Summary of word order patterns

Before turning to the theoretical discussion of the above data, let us take stock of the findings so far. As we have seen, the exclamative phrase in *de*-exclamatives and *wh*-exclamatives is always left-adjacent to the verb. Following the E-phrase, the order of verb and preverb is variable with some E-expressions but not with others. The variation in this domain is summarized in Tables 1 and 2. Table 1 lists *wh*-phrases which do not contain other lexical material that might influence the distribution of the E-phrase. Table 2 lists the latter type of E-phrases separately: that in which *milyen/de* 'what.kind, how' modifies a quantifier or an adverbial:

Table 1

The distribution of *wh*- and *de*-phrases

	inversion	straight order
mennyire, hogy 'to what extent/how much'	*	✓
hány, mennyi 'how much/many'	✓	✓
plurals (*kik* 'who-pl', *mik* 'what-pl')	✓	?
quantified phrases (*ki mindenki*'who all')	%	✓
ki 'who', *mi* 'what', *mikor* 'when', *hol* 'where' *hogy(an)* 'how', *melyik (N)* 'which (N)', *milyen/de* Adj N 'how Adj N', etc.	✓	*

Table 2

The distribution of *de/milyen*...-phrases

		inversion	straight order
with quantifiers	*sok (N)* 'many/much'	✓	✓
	kevés (N) 'few/little'	✓	*
with adverb type	grade, completion, intensity (*nagyon* 'very much')	*	✓
	durván ('in a rough way')-class	✓	✓
	exclusive (*késõn* 'late')	✓	*
	lassan ('slowly')-class	✓	*

It is important to stress again that this classification is not fully exhaustive. Rather, its aim is to highlight types of phrases whose behaviour is characteristic for a particular class of items.

3. The syntactic structure of Hungarian exclamatives

3.1. An initial comparison with interrogatives

Exclamative sentences differ from declaratives and interrogatives functionally: instead of asserting something (as do declaratives) or questioning something (as do interrogatives), they express surprise or astonishment about something outstanding. How is this function coded in Hungarian? Next to lexical marking (the *de* marker) and special phonology, is there a special syntactic structure which exclamatives assume?

The most instructive manner of looking at this question is to compare exclamatives to interrogatives. Such comparison suggests itself because in many languages there is a striking similarity between exclamatives and questions: they can both contain *wh*-phrases, and they share a number of significant properties, like the fact that *wh*-movement into the initial position in the sentence is obligatory in both:

(31) (a) [What book] did you buy?

(b) [What a strange book] you bought!

The presence of such structural parallels in some languages has given rise to analyses that try to fully derive exclamatives from questions or give a similar account for both (D'Avis 2002; Pesetsky Torrego 2001; Fujii Ono 2005). At the same time, there also exist languages in which placement of exclamative *wh*-phrases does not follow the placement of *wh*-phrases in

questions. French (Obenauer 1976) and Italian (dialects) (Munaro 2003; Portner–Zanuttini 2003) are known examples. In Paduan, for example, *wh*-constituents in questions always follow left dislocations (*a to sorela* 'to your sister'), while complex *wh*-phrases in exclamatives precede them (Portner–Zanuttini 2003):

(32) (a) A to sorela, [che libro] vorissi-to regalar-ghe? question
to your sister which book want-cl give-her
'To your sister, which book would you like to give as a gift?'

(b) [Che bel libro], a to sorela, che i ghe ga regalà! exclamative
what nice book to your sister that cl her have given
'What a nice book, to your sister, they gave her as a gift!'

The situation in Hungarian exclamatives recalls the state of affairs from Paduan: while interrogative *wh*-phrases without exception get fronted into the immediately preverbal position triggering inversion between the verb and the preverb, some exclamative *wh*-phrases can occupy a position which seems to be distinct in the light of inversion.[8] Inversion can be missing sometimes with some E-phrases, as was noted above:

(33) (a) Hány filmet {*megnéztél / néztél meg}?
how.many film-acc pv-watched-2sg watched-2sg pv
'How many films did you watch?'

(b) (Hogy) hány filmet {megnéztél / néztél meg}!
comp how.many film-acc pv-watched-2sg watched-2sg pv
'You watched so many films!'

Research on Hungarian (Horvath 1981; É. Kiss 1987; Brody 1995) has repeatedly analyzed inversion as indicative of some constituent moving to the unique contrastive focus position. Lack of inversion on the other hand is indicative of a distinct configuration: one in which there is no focused element in the sentence. From this it follows that the position of the exclamative phrase in (33b) is not the same as in (33a). And this in turn rules out the possibility of an analysis that would claim that the derivation of exclamatives runs fully parallel to questions. As will be shown in the rest of the paper, the derivations of the two are not the

[8] As the attentive reader will have noticed, lack of subject auxiliary inversion also characterizes exclamatives in English, cf. (31) above. See Pesetsky–Torrego (2001), Fujii–Ono (2005) for an account of this.

same. With some E-phrases the derivation can partially overlap with that of questions, but never in a fully identical manner.

To see this precisely, the next two sections proceed to elaborate on the precise placement of E-phrases. Section **3.2** deals with the items in Table 2, and section **3.3** turns to Table 1.

3.2. The placement of adverbial and quantificational E-phrases (Table 2)

The distribution of E-phrases listed in Table 2 can easily be described by a simple principle, as these phrases show strict regularities when compared to their behaviour in indicative, neutral clauses. They occupy the same position in exclamatives and in indicatives. Consider the following two illustrative examples, repeated from above:

(34) (a) De/milyen alaposan {megfázott / *fázott meg} Ágnes!
de how thoroughly pv-cold.caught-3sg cold.caught-3sg pv Ágnes
'What a thorough cold Ágnes got!'

(b) Ágnes alaposan {megfázott / *fázott meg}.
Ágnes thoroughly pv-cold.caught-3sg cold.caught-3sg pv
'Ágnes caught a thorough cold.'

(35) (a) {De / milyen} kevés könyvet {olvastál el / *elolvastál}!
de how few book-acc read-2sg pv pv-read-2sg
'How few books you read!'

(b) Kevés könyvet {olvastál el / *elolvastál}.
few book-acc read-2sg pv pv-read-2sg
'You read few books.'

As (34b) shows, the intensity adverb *alaposan* 'thoroughly' can never occur with inversion in declarative sentences. The reason for such a characteristic positioning is presumably lexical: it has to do with the meaning of the adverbial (its features such as +grade, +contrast, Kiefer 1967). Phrases with *kevés* 'few/little' on the other hand show the opposite behaviour: they can only occur with inversion in neutral sentences, due to the lexical semantics of *kevés* 'few' (Szabolcsi 1997). Importantly, both *alaposan* and *kevés* keep their behaviour in exclamatives as well, leading to the generalization in (36):

(36) The position of adverbial E-phrases in exclamatives is the same as the position they occupy in indicatives when they carry the main stress of the sentence.

The relevance of stress might not be clear at first sight, so it deserves specific attention. The generalization in (36) states that the positions the E-phrase occupies in exclamatives correspond to stressed positions in indicatives. This effect can best be observed in the behaviour of *lassan*-type adverbials. As we have shown above in (29) and (30b), repeated here as (37a) and (37b), *lassan*-type adverbs in exclamatives occur with the word order they exhibit in indicatives when they are stressed, but crucially not with the word order they exhibit in indicatives when they are not stressed:

(37) (a) {De / milyen} lassan {érett meg / *megérett} a szilva!
de how slowly ripened-3sg pv pv-ripened-3sg the plum
'The plums got ripe so slowly!'

(b) 'Lassan {érett meg / *megérett} a szilva.
slowly ripened-3sg pv pv-ripened-3sg the plum
'The plums underwent slow ripening. (It is not possible that the ripening itself went quickly.)'

Unstressed *lassan* occurs without inversion and with a different meaning, and this placement (as well as reading) is ruled out in exclamatives (cf. 37a):

(38) Lassan megérett a szilva.
slowly pv-ripened-3sg the plum
'Slowly, the plums got ripe. (It is possible that the ripening itself went quickly.)'

(36) thus captures the correspondence between exclamative placement and stressed positions in the left periphery. (36) also covers other adverbial phrases or quantificational phrases with *sok* 'many, much' and *kevés* 'few, little'. As the reader can check for himself, all these phrases are stressed in their preverbal position they occupy in neutral indicative clauses.

Before closing this section, a note is in order about the topic of the previous section, the comparison between exclamatives and interrogatives. This is necessary because the findings in (36) have interesting repercussions for this topic as well. Unlike exclamatives, interrogative sentences leave no room for optionality in the placement of *wh*-phrases (33a): they are uniformly placed in the syntax, triggering inversion. Movement

to this position is usually taken to be triggered by the *wh*-feature on the *wh*-phrase (Chomsky 1995). The fact that exclamative sentences do not force the distribution of question-phrases onto E-phrases can be due to two reasons. Either exclamative syntax is "blind" to the presence of *wh*-features, or these features are not present on *wh*-E-phrases to begin with. In the light of previous work that claims that *wh*-words do not have an inherently interrogative meaning (Lipták 2001, which follows Cheng 1991), I contend that *wh*-words in exclamatives have no *wh*-feature, which explains why they do not have interrogative word order. The *wh*-feature that drives movement in interrogatives only characterizes *wh*-words in interrogatives.[9] Exclamative *wh*-phrases for example (similarly to relative pronouns, or indefinite *wh*-items, see Lipták *op.cit.*, chapter 4) do not carry *wh*-features that drive movement in interrogatives.

3.3. The distribution of other wh-phrases in exclamatives (Table 1)

While the distribution of the items in Table 2 was easy to account for, Table 1 is much more difficult to capture in a simple generalization. It is not surprising of course: since *wh*-phrases do not occur in indicative clauses, there is nothing to compare exclamatives to, in the indicative domain. Comparison with interrogatives does not reveal full parallels, either, since the distribution of *wh*-phrases in interrogatives is uniform and exceptionless: they all trigger inversion.

At the same time, the distribution of E-phrases in Table 1 does show certain regularities that recall parallels with indicative sentences. Before turning to these syntactic observations, the next subsection introduces the layout of the Hungarian left periphery in detail to prepare the ground.

3.3.1. The structure of the Hungarian left periphery

Phrases that occur adjacent to the left of inverted verbs are standardly analyzed as contrastive **focus**, occupying the specifier position of FocP,

[9] In interrogatives, *wh*-expressions are bound by a word-level question operator morpheme (Q_{wh}), which provides them with question semantics, and carries the feature that drives overt movement of *wh*-phrases to FocP. For further details see Lipták (2001). In non-interrogatives, no such Q_{wh} operator is present.

a distinguished position for such constituents.[10] When the specifier of FocP hosts a focused phrase (lexical focus or interrogative *wh*-phrase), the head of FocP has to be filled by the verb. This triggers obligatory inversion between the verb and the preverb, if the latter is present: the verb strands its preverb in a position lower than FocP (possibly in AspP). Focusing and verb raising to F^0 has the fine structure illustrated in (39):

(39) $[_{\text{FocP}}$ {focus} $[_{\text{Foc}'}$ V^0_i ... $[_{\text{AspP}}$ pv ... $[_{\text{VP}}$ t_i]]]]

FocP is a rather low projection in the Hungarian left periphery. It is dominated by a set of other left peripheral projections, most importantly the functional projections hosting distributive quantifiers (DistPs), topics (TopPs) and the complementizer projection (CP). These projections are ordered in the following way:

(40) $[_{\text{CP}}$ $[_{\text{TopP*}}$ $[_{\text{DistP*}}$ $[_{\text{FocP}}$ {focus} V^0 $[_{\text{AspP}}$ pv ...]]]]]

Following the complementizer and topics, DistP is the projection of universal quantifiers in the left periphery. This projection was termed QP in É. Kiss (1987), and later came to be known as a distributive projection (DistP) due to Szabolcsi (1997), who argued that this position is unique in only hosting distributive constituents.

That DistP is a projection distinct from FocP can be seen from the fact that universal quantifiers (i) cannot occur in Spec,FocP themselves, i.e., do not trigger inversion and (ii) are not in complementary distribution with a focused expression. These properties are illustrated in (41) and (42). (41) shows that a universal quantifier does not trigger inversion and is not compatible with it:

(41) (a) Mindenkit meghívott János az ünnepségre.
everyone-acc pv-invited-3sg János the celebration-on
'János invited everyone to the celebration.'

[10] The structural positions to be reviewed here reflect the result of "standard" research that is most often adopted for simple analyses of the Hungarian left periphery. It is important to keep in mind that this model has recently been criticized and newest developments have questioned the existence of FocP, and verb movement to Foc^0 (Horvath 2000; Koopman–Szabolcsi 2000; Olsvay 2000, to mention some). Such developments do not affect the current discussion in critical ways.

(b) *MINDENKIT hívott meg János az ünnepségre.
everyone-acc invited-3sg pv János the celebration-on
intended: 'It was everyone whom János invited to the celebration.'

(42) illustrates that universal quantifiers only allow a focus or a verb to follow them:

(42) (a) Mindenkit JÁNOS hívott meg az ünnepségre.
everyone-acc János invited-3sg pv the celebration-on
'It was János who invited everyone to the celebration.'

(b) Mindenkit (*tegnap) meghívott az ünnepségre János.
everyone-acc yesterday pv-invited-3sg the celebration-on János
'Yesterday János invited everyone to the celebration.'

This is in accordance with the structure in (40) which registers the fact that DistP dominates FocP in Hungarian.

Recent work (Kálmán 2001) has argued that the DistP projection should rather be characterized as a DistP **field** comprising several slightly distinct projections. The split of the DistP projection is most notably required by the empirical properties of emphatic *sok* 'many/much'-phrases, which also occupy a DistP position when in the left periphery ("'"stands for emphasis):[11]

(43) 'Sok lányt meghívott János az ünnepségre.
many girl-acc pv-invited-3sg János the celebration-on
'János invited many girls to the celebration.'

Emphatic *sok*-phrases express the speaker's judgement about a high amount or numeric degree. (43), for example, indicates that according to the speaker's judgement, there were many invited girls (above average, above expectation or contrasting with only a few girls).

Evaluative *sok*-phrases are different from universal quantifiers in two respects. One is that they can occur in Spec,FocP as focused constituents (compare this with (41b) above):

(44) SOK LÁNYT hívott meg János az ünnepségre.
many girl-acc invited-3sg pv János the celebration-on
'It was many girls who János invited to the celebration.'

[11] *Sok*-phrases without emphasis have a wider distribution. They can occur as topics or postverbal constituents as well.

The other is that emphatic *sok*-expressions always follow but do not precede universal quantifiers when the latter are also present in the left periphery:

(45) (a) Mindenhova 'sok lányt meghívott János.
everywhere many girl-acc pv-invited-3sg János
'János invited many girls to every place.'

(b) *'Sok lányt mindenhova meghívott János.
many girl-acc everywhere pv-invited-3sg János
'János invited many girls to every place.'

Universal quantifiers on the other hand have no ordering restrictions among themselves:

(46) (a) Mindenkit mindenhova meghívott János.
everyone-acc everywhere pv-invited János

(b) Mindenhova mindenkit meghívott János.
everywhere everyone-acc pv-invited János
'János invited everyone to every place.'

To accommodate the observed co-occurrence restrictions of quantificational phrases, one needs to assume the structure in (47): a unique *many*P for *sok*-expressions, which is distinct from DistP.[12] According to the testimony of the facts above, emphatic *sok*-phrases can occupy either this *many*P (cf. 47a) or that of focus (47b).

(47) (a) [... [$_{manyP}$ *sok*-XP [$_{AspP}$ pv-V [...]]]]

(b) [... [$_{FocP}$ *sok*-XP V^0 [$_{AspP}$ pv [...]]]

When the *sok*-phrase occupies Spec, *manyP*, it is adjacent to a non-inverted pv-V verb, which I assume stays in AspP, right below *many*P.[13] When the *sok*-phrase occupies the focus position, it forces verb preverb

[12] The anonymous reviewer calls my attention to the fact that not all speakers seem to make the grammatical distinction between evaluative *sok*-phrases and universal quantifiers. For these speakers (45b) is grammatical. I consulted four extra speakers to check this point, and indeed one speaker out of the four allows for (45b). I refer such individual variation to future research.

[13] The adjacency requirement between the *sok*-phrase and the verb (similarly to that in (42b)) has not yet received explanation in the literature on Hungarian to my knowledge. Putting it down to the selectional restriction of *many*P suffices for the purposes of this paper but it is nothing more than a mere descriptive coding

inversion, as focus constituents do. The structures in (47) will form the basis of the discussion in the next sections. The first section will analyze *wh*-E-phrases that can occur both with and without inversion. The second one deals with those that can only occur with inversion. The third one treats those which can only occur without inversion.

3.3.2. The position of *wh*-E-phrases with optional inversion

Let us start the discussion with the most characteristic type of *wh*-phrases that occur with optional inversion *wh*-phrases: *hány* and *mennyi* 'how many/much', which denote amount. These phrases show the exact same syntactic behaviour as the above mentioned *sok* 'many, much' phrases in indicatives:

(48) (a) (Hogy) hány könyvet {megvettél / vettél meg}!
comp how.many book-acc pv-bought-2sg bought-2sg pv
'You bought so many books!'

(b) 'Sok könyvet {megvettél / vettél meg}.
many book-acc pv-bought-2sg bought-2sg pv
'You bought many books.'

This parallel suggests that the placement of *hány/mennyi* in (48a) and that of the *sok*-phrase in (48b) is identical: when the phrase occurs with inversion, it is focused in Spec,FocP, and when it occurs without inversion, it occupies *manyP* (cf. 47) (for a similar suggestion, see Kálmán 2001). This state of affairs can also be supported by other parallels between evaluative *sok*-phrases and *hány/mennyi*-phrases in exclamatives.

One such parallel is the fact that *sok*-phrases and *hány/mennyi*-phrases are adjacent to the pv-V sequence (when they occur without inversion):

(49) (a) (Hogy) hány könyvet (*tegnap) megvettél!
comp how.many book-acc yesterday pv-bought-2sg
'You bought so many books yesterday!'

(b) 'Sok lányt (*tegnap) meghívott János az ünnepségre.
many girl-acc yesterday pv-invited-3sg János the celebration-on
'János invited many girls to the celebration.'

---

of the observed facts. The real explanation behind this adjacency presumably lies elsewhere.

Another parallel between the two concerns co-occurrence restrictions with lexical focused phrases. It seems that some speakers disallow a lexical focus after an evaluative *sok*-phrase (cf. 50a), similarly to the varying judgements (all) speakers provide for cases in which a *hány/mennyi*-phrase precedes focus (cf. 50b):

(50) (a) %'Sok lányt JÁNOS hívott meg az ünnepségre.
many girl-acc János invited pv the celebration-on
intended: 'It was János who invited many girls to the celebration.'

(b) %(Hogy) hány könyvet JÁNOS vett meg!
comp how.many book-acc János bought-3sg pv
'How many books JÁNOS bought!'

This parallel between the two types of constructions is arguably less strong due to the fact that both are subject to substantial individual variation, details of which are not completely clear to me. The fact, however, that both sentence types are in any event clearly marked does not run counter to the claim that the position of *sok-* and *hány*-phrases in them can be similar.

The above parallels single out one possible structural position that *hány/mennyi* can occupy: *many*P. Universal quantifiers, which inhabit DistP, do not show the pattern in (50): they allow for a focus following them without any problem (see (42a) above). The same is also true about even higher left peripheral constituents, like topics. They do not only differ from exclamatives in the property in (50), but also in the property in (49): they need not be adjacent to a verb and can be followed by focus as well as other quantifiers or topics.

This identifies the position of *hány/mennyi* E-phrases as that of emphatic *many*P, which, as argued before, is a position distinct from that of focus (51a). When these items occur with inversion, they occupy the focus position (51b):

(51) (a) [... [*many*P {*hány/mennyi*} [AspP pv-V [...]]]]
(b) [... [FocP {*hány/mennyi*} V^0 [AspP pv ...]]]

The distinct syntactic placement in the two cases has an effect on both semantic properties and syntactic behaviour, further supporting the structures in (51). Two important properties need mention here: distributive vs. collective readings and the licensing of postverbal superlatives.

As noted above, *many*P is part of the quantificational field, the DistP field of Hungarian. Constituents in the DistP field have an obligatorily distributive reading. This is in stark contrast with FocP, which can host constituents with both distributive and collective readings (Szabolcsi 1997). Due to this essential difference, the meaning of *sok*-phrases differs in distributivity depending on their structural position. When they are in *many*P, i.e., in the quantificational field, not triggering inversion, they are obligatorily distributive. When they are in FocP, triggering inversion, they are optionally distributive:

(52) (a) 'Sok gyerek felemelte a zongorát.
many kid pv-lifted-3sg the piano-acc
'Many kids lifted the piano (separately).'

(b) SOK GYEREK emelte fel a zongorát.
many kid lifted-3sg pv the piano-acc
'Many kids lifted the piano (separately/together) (not just a few).'

The exact same phenomenon can be observed with *hány/mennyi*-phrases that are allowed to appear in both positions:

(53) (a) (Hogy) hány gyerek felemelte a zongorát!
comp how.many kid pv-lifted-3sg the piano-acc
'How many kids lifted the piano (separately)!'

(b) (Hogy) hány gyerek emelte fel a zongorát!
comp how.many kid lifted-3sg pv the piano-acc
'How many kids lifted the piano (separately/together)!'

The positional difference sketched in (51) results in syntactic differences between the two patterns, too. One such difference concerns the licensing of postverbal superlative expressions. The licensing of superlatives can only be done from the focus position, and not from the quantificational position, as was argued in É. Kiss–Farkas (2001):

(54) (a) JÁNOS itta meg a legkevesebb bort.
János drank-3sg pv the least wine-acc
'It was János who drank the least wine.'

(b) *János minden nap megitta a legkevesebb bort.
János every day pv-drank-3sg the least wine-acc
'János drank the least wine every day.'

Parallel to the facts in (54), *sok*-phrases license a superlative when they are syntactically focused in Spec,FocP, but not when they are in *manyP*, as (55) shows:

(55) (a) János sokszor ért ide a leggyorsabban.
János often arrived-3sg pv the quickest

(b) *János sokszor ideért a leggyorsabban.
János often pv-arrived-3sg the quickest
'János often arrived here the most quickly.'

The same is true about exclamatives when they appear in different positions. Exclamative *wh*-phrases can only license a superlative phrase from the focus position (i.e., with inversion), but not from *many*P (without inversion):

(56) (a) (Hogy) hányszor értél ide a leggyorsabban!
comp how.often arrived-2sg pv the quickest

(b) *(Hogy) hányszor ideértél a leggyorsabban!
comp how.often pv-arrived-2sg the quickest
'How often did you arrive here the most quickly!'

On the basis of these facts, there remains little doubt that the structures in (51) (parallel to those in (47)) are on the right track about exclamatives: just like emphatic *sok*-phrases, exclamative *hány/mennyi*-phrases occupy either Spec,*many*P or Spec,FocP. Does this come as a surprise? Certainly not, since the two types of phrases share common semantic features. Both emphatic *sok*-phrases and exclamative *hány/menynyi*-phrases are **evaluative** expressions (expressing the speaker's judgement) with a high **amount** reading. *Many*P in Hungarian subcategorizes for constituents with these two features. This is what allows for *hány/mennyi* to occur in this position in exclamatives. Note furthermore that *many*P seems to be only compatible with **high** amount readings. While *sok*-phrases express **high** amount as part of their lexical meaning, *hány/mennyi*-phrases do not. They are in principle compatible with both high and low amount readings. Placement in *many*P, however, singles out the high amount reading (Kálmán 2001), strengthening the claim that the *hány*-expression is in *many*P:

(57) (a) (Hogy) hány filmet megnéztél!
comp how.many film-acc pv-watched-2sg
'You watched so many films! / *You watched so few films!'

(b) (Hogy) hány filmet néztél meg!
comp how.many film-acc watched-2sg pv
'You watched so many films! / You watched so few films!'

Having accounted for *hány/mennyi*, what is left now is to account for quantified E-phrases and plurals, which show a similar behaviour to *hány/mennyi*:

(58) (a) (Hogy) ki mindenki {eljött / ?jött el} az ünnepségre! (= (22))
comp who everyone pv-came-3sg came-3sg pv the celebration-to
'The (different) kinds of people/the number of people who came to the celebration!'

(b) (Hogy) mi mindent {megettél / ?ettél meg}! (= (23))
comp what everything-acc pv-ate-2sg ate-2sg pv
'The number of things you have eaten!'

I contend that quantified *wh*-phrases, like *ki mindenki* 'who all' and plurals like *kik* 'who-pl' also have a similar optionality in placement as *sok*-phrases and *hány/mennyi*-phrases described above. This is due to the fact that these phrases are **quantificational** phrases, expressing high amount as well.

That E-phrases with the universal quantifier *minden* 'every' express high amount is beyond doubt. The presence of the quantifier in these phrases is presumably linked to the fact that exclamatives are scalar expressions (see section 4).[14] They always invoke a scale on which the E-phrase denotes a high scalar value. It is likely that the presence of the universal quantifier emphasizes some property of this scale: for example that the individuals range over various values of the scale. Leaving the implementation of this intuition aside, the amount reading of quantified E-phrases can be accounted for along these lines. The fact that these *minden* 'every'-phrases are slightly dispreferred in the focus position sug-

[14] The exclamative nature of such universal quantifiers is clearly observable in languages in which these elements are strongly required to indicate exclamative use of the sentence they occur in. Dutch *allemaal* 'all', (although not brilliant in the example in (i)) is necessary to make the exclamative interpretation salient (Marcel den Dikken, p.c.):

(i) Wat je ??(?allemaal) moet doen om aan een baan te komen! (Dutch)
what you all must do to on a job to come
'The things you have to do to get a job!'

gests furthermore that the quantificational nature of these phrases is so strong that their default placement is in the quantificational domain.

Plural E-phrases in my opinion need to receive a similar analysis: they can, at least for some speakers, appear in the *many*P position. This is presumably linked to the fact that plurality can be conceived of as quantification in semantics (Link 1983). For speakers who assign quantificational value to these plural phrases, *many*P is an acceptable position. For speakers who do not treat these as quantificational, they pattern with ordinary *ki* 'who', *mi* 'what' *wh*-phrases that always occur in the focus position.

Summing up, this section provided arguments to the effect that E-phrases that can occur both with and without inversion can occupy two positions in the Hungarian left periphery, the specifier of *many*P or that of FocP:

(59) (a) [... [$_{manyP}$ {*hány/mennyi / ki mindenki / %kik*} [$_{AspP}$ pv-V [...]]]]
(b) [... [$_{FocP}$ {*hány/mennyi / ?ki mindenki / kik*} V^0 [$_{AspP}$ pv ...]]]

These two positions were identified and told apart using evidence from (i) the distribution of verb movement (inversion), (ii) co-occurrence restrictions of E-phrases with other constituents, (iii) licensing of superlative postverbal constituents, (iv) the availability of distributive/collective readings, and (v) the availability of high and low amount readings. Taking the lead of this structural analysis, it was concluded that the class of *wh*-E-phrases with the distribution in (59) comprise high amount expressions with an evaluative reading. This is because *many*P is selective as to the kind of constituents it hosts: it only allows for amount expressions with evaluative meanings. While evaluative meaning characterizes all exclamative expressions, only *hány/mennyi, ki mindenki* and *kik*-type expressions express high amount lexically. This is why their placement allows for the kind of optionality observed.

The distribution of other types of *wh*-phrases will be discussed in the next section.

3.3.3. The position of *wh*-E-phrases with obligatory inversion and obligatory straight order

With the analysis of *hány/mennyi*-expressions in place, identifying the position of E-phrases with an obligatory placement becomes significantly easier.

Starting the discussion with those *wh*-E-phrases that occur with obligatory inversion (those that denote individuals, times, manners and reasons like *ki* 'who', *mi* 'what', *hol* 'where', *mikor* 'when', *hogyan* 'how' etc.) the picture is very clear. As inversion is the key characteristic of Spec,FocP being filled in Hungarian, the conclusion is that these elements are in FocP:

(60) [... [$_{\text{FocP}}$ {*ki / mi / hol / mikor / hogyan*} V^0 [$_{\text{AspP}}$ pv ...]]]

The focal placement of these items is further supported by the tests used in the previous section. E-phrases in this position are optionally distributive or collective (61), and they license a postverbal superlative expression (61b):

(61) (a) (Hogy) melyik két gyerek emelte fel a zongorát!
comp which two kid lifted-3sg pv the piano-acc
'(I am surprised at) which two kids lifted the piano (separately/together)!'

(b) (Hogy) mikor értél ide a leggyorsabban!
comp when reached-2sg pv the quickest
'(I am surprised at) the time you got here as the quickest!'

While the focal placement of *ki* 'who', *mi* 'what', *hol* 'where', *mikor* 'when', *hogyan* 'how'-type phrases is beyond doubt, the question why they receive such a placement is more difficult to answer. Upon first sight, one is inclined to connect this property to the *wh*-hood of these constituents: after all, *wh*-phrases also occupy the focus position in interrogatives. If *wh*-phrases in exclamatives are like *wh*-phrases in interrogatives, a similar distribution is expected.

This line of reasoning, however, cannot be on the right track for various reasons. The most robust of these is that obligatory focusing does not only affect *wh*-phrases. Some *de*-phrases also trigger obligatory inversion, as was illustrated in (26d) above:

(62) De drága könyvet {vettél meg / *megvettél}! (= (26d))
de expensive book-acc bought-2sg pv pv-bought-2sg
'How expensive a book you bought!'

The behaviour of such *de*-phrases rules out the possibility of linking focal placement to *wh*-hood.

Other reasons not to connect focal behaviour with *wh*-hood involve the lack of full parallels between exclamatives and interrogatives, as was

mentioned already. To repeat these, the syntax of exclamatives was shown not to be parallel to that of interrogatives in that (i) exclamatives, but not interrogatives, can involve non-*wh*-phrases (*de*-phrases, cf. (2)), (ii) exclamatives allow for some *wh*-phrases in positions that are not available to *wh*-phrases in interrogatives (section **3.1**), (iii) matrix exclamatives, unlike matrix interrogatives can have an overt complementizer (cf. (4)).

Due to these considerations the focal placement of *ki, mi, hol, mikor, hogyan*-type phrases cannot be due to a parallel with interrogatives. The Hungarian facts are not compatible with analyses that try to reduce exclamatives to interrogatives (D'Avis 2002) and claim that the two clause types have identical syntax when it comes to the placement of *wh*-phrases (as is done for English in Pesetsky–Torrego 2001 or Fujii–Ono 2005).

What drives the movement of *ki, mi, hol, mikor, hogyan*-type phrases then? It is clear that whatever it is, it must be distinct from the driving force of interrogative *wh*-phrases. The property that drives this movement is likely to be inherent to every exclamative phrase, not only those confined to Spec,FocP, namely the *ki, mi, hol, mikor, hogyan*-type. The unifying property of all E-phrases is that they are evaluative scalar expressions. I propose that focus placement of E-phrases follows from this very property: being scalar requires focusing. It is the evaluative scalar nature of *ki, mi, hol, mikor, hogyan*-type elements that allows and forces them to occur in the focus position. Section 4 below will elaborate on this idea in more detail, showing how scalarity affects the form and placement of exclamative phrases.

3.3.4. The position of *wh*-E-phrases with obligatory straight order

Before closing this section, however, the distribution of yet another type of E-phrases need to be accounted for: those that never occur without inversion. It seems we can be short about these items, as their behaviour is rather exceptional and almost idiomatic. There are two *wh*-E-phrases with this property: *mennyire* and *hogy* 'to what extent/how much' denoting extent. Following the logic of the findings in the previous section, the fact that these never occur with inversion shows that they are not in FocP. Instead, *manyP* suggests itself as a possible candidate. It seems, however, that the position of the *mennyire*-phrase is not in *many*P, but rather the position that completion/intensity adverbs like *nagyon* 'very much' or *alaposan* 'thoroughly' also occupy. These kinds of adverbials

also always occur without inversion and express the exact same meaning of extent:

(63) (a) Éva nagyon {megnőtt / *nőtt meg}.
Éva much pv-grew-3sg grew-3sg pv
'Éva has grown a lot.'

(b) *Nagyon ÉVA nőtt meg.
much Éva grew-3sg pv
'It was Éva who has grown a lot.'

Since *mennyire*-phrases also express high extent or intensity (rather than amount, as *sok*-phrases or *hány/mennyi* E-phrases), I take it that they assume the placement of *nagyon*-type adverbials:

(64) [... [AdvP {*nagyon / mennyire*} [AspP pv-V [...]]]

It needs to be mentioned that the position of *nagyon*-type grade, completion or intensity adverbials is indistinguishable from that of emphatic *sok*-phrases syntactically: (i) *sok*- and *nagyon*-phrases are in complementary distribution, (ii) both types can only stand without inversion and (iii) they do not tolerate a focused constituent to their right. Leaving a comparison between *nagyon*- and *sok*-phrases for further research, the rest of the paper will not deal with this kind of data any more.

4. Focus in the theory of exclamatives

The previous section gave a structural analysis of E-phrases in Hungarian, and delivered the result that the placement of exclamatives involves two possible positions: *many*P and FocP:

(65) (a) [... [*many*P {*hány / mennyi / ki mindenki / %kik*} [AspP pv-V [...]]]]
(b) [... [FocP {*hány / mennyi / ?ki mindenki / kik*} V^0 [AspP pv ...]]]
(c) [... [FocP {*ki / mi / hol / mikor / hogyan*} V^0 [AspP pv ...]]]

Due to lexical restrictions of the former, only amount expressions fit into *many*P, while all exclamative phrases can occupy the focus slot (with the exception of extent-denoting ones in **3.3.4**, put aside as idiomatic ones). The present section aims at explaining this distribution, concentrating primarily on the role focusing plays in exclamative syntax. The first subsection will spell out the semantics of exclamations briefly. The second

subsection will point out that focusing in the syntax is a natural consequence of exclamative semantics. This section will also provide some analysis of relativized exclamatives (cf. example (3)).

4.1. The semantic characteristics of exclamatives

The semantics of exclamatives has been reasonably well-studied in recent years. The following list of semantic properties is based on pioneering work in Elliott (1974), Grimshaw (1979), Michaelis–Lambrecht (1996) and Portner–Zanuttini (2003):

(66) Semantic properties of exclamatives

(a) FACTIVITY. The proposition exclamatives denote is entailed.

(b) SCALAR IMPLICATURE. Exclamatives assert that the degree of a particular scalar property lies at the extreme end of a (contextually given) scale.

(c) INVOKING A SET. Exclamatives invoke a set of alternative propositions; and they indicate that the proposition in this set exceeds the expected ones (the latter called "widening" in Portner–Zanuttini 2003).

(d) EXPECTATION CONTRAVENTION. Exclamatives express that something is unusual and surprising.

(e) EMOTIVENESS. Exclamatives assert an affective stance towards a propositional statement.

Factivity is a major semantic property of exclamatives. Exclamatives introduce the presupposition that their propositional content is true. When uttering (67), the speaker subscribes to the fact that the addressee watched some films:

(67) (Hogy) hány filmet megnéztél!
comp how.many film-acc pv-watched-2sg
'You watched so many films!'

Due to the factive property, exclamatives cannot be embedded under non-factive predicates (see also (6)–(8) in section **2** above):

(68) *Úgy tudom, hogy hány filmet megnéztél.
so know-1sg comp how.many film-acc pv-watched-2sg
'I know that you watched many films.'

Next to factivity, the scalar and set-invoking nature of exclamatives is also considered to be defining properties of exclamations in Portner–

Zanuttini (2003). Scalarity refers to the fact that exclamatives always operate on a scale: the surprising property that the exclamative expresses is placed on a scale that contains alternative values corresponding to various degrees, ranging from small to high degrees. The exclamative singles out an extreme degree on this scale.

It is important to note that scalarity characterizes all exclamatives across the board, irrespective of whether they contain lexical elements that can be associated with a scale. Phrases of gradable adjectives, for example, have been claimed to include a specification of degree (Corver 1990). This makes E-phrases with adjectives perfect exclamative phrases. Yet, E-phrases of other categories also get associated with a scale in Hungarian. Nominals for example, which are standardly not associated with degrees, get associated with a high degree property in exclamatives. (69) is associated with a scale of importance, and indicates that the visiting person is exceptional in this respect:

(69) (Hogy) ki jött el ebbe a faluba!
comp who came-3sg pv this-into the village-into
'What a person came to this village!'

The set-invoking property of exclamatives (property iii) means that exclamatives invoke a set of **alternative propositions**; the use of exclamatives is to point out that the true proposition among these alternatives is the one that exceeds the expected one. In this sense, exclamatives **widen** the original scale on which they operate. Widening means that the extreme value they define falls outside the standard scale. The sentence in (67), for example, indicates that the number of films the addressee watched is greater than the alternatives under consideration, namely being one film, or a few films, or an expected number of films. (69) indicates that the importance of the visiting person is higher than normal.

Widening thus gives rise to yet another important property of exclamatives: expectation contravention. This means that exclamatives express that something is unexpected, unusual, out of the ordinary. The exclamative lexicon sometimes provides first-hand evidence for the expectation contravention nature of exclamatives. Functional heads or exclamative morphemes in exclamatives often originate from words with an emotive or adversative meaning. A telling example is the adversative

coordinator *de* 'but' in Hungarian, which is the exclamative word used in what I refer to as *de*-exclamatives:[15]

(70) *De* jó volt az a buli!
but good was that the party
'How great that party was!'

Exclamative *de* arguably originates from the adversative coordinator 'but', whose role as a coordinator is to express the contrast by indicating expectation contravention.

The last key characteristic of exclamatives is emotiveness. Exclamatives express the speaker's surprise, which is an affective (emotive) stance. Emotivity is responsible for the fact that exclamatives often contain expressive lexical items (on the latter in general see Potts–Roeper 2006), as was observed above in (9a):

(71) Milyen rohadtul megfáztam!
how rottenly pv-cold.caught-1sg
'What an awful cold I got!'

In another context the emotive nature of exclamatives is lexicalized by **csoda** 'wonder', an emotive word (den Dikken–Lipták 1997), which sometimes forms obligatory part of the exclamation:

(72) (a) Mi-*(csoda) egy fickó!
what-wonder a guy
'What a guy!'

(b) Mi a csoda!
what the wonder
[idiomatic ejection of surprise]

This concludes the description of the semantic characteristics of exclamatives. Of these characteristics two prove to be relevant for the syntactic analysis of exclamatives that is to be implemented in this paper: scalar

[15] Serbo-Croation has a similar use of the adversative coordinantor (Boban Arsenijević, p. c.):

(i) *Al* je Jovan snazan!
but is Jovan strong
'How strong is Jovan!'

implicature and invoking a set. The next section will show how these properties account for the syntactic placement of E-phrases in Hungarian.

4.2. The role of focus in Hungarian exclamatives

The question that section **3** ended upon was (recall (65a, b, c)): why can the syntactic FocP position host E-phrases in Hungarian and why does it **have** to host the kinds that cannot occur in *many*P? In the light of the observed semantic properties of exclamatives, this question can now be answered in the following way.

The syntactic focus placement of exclamative phrases follows from the fact that exclamatives invoke a set of alternatives. As was shown above, exclamatives always invoke a set of **alternative propositions**. The role of this set of alternative propositions is to spell out what the exclamative proposition differs from: the exclamative states that the true proposition among these alternatives exceeds the expected ones. Now, recalling a set of alternatives characterizes contrastive focus constructions in general across languages, including the Hungarian contrastive focus placed in Spec,FocP (Kenesei 1986; Rooth 1992).[16] For illustration, consider the following example with a lexical focus in FocP:

(73) A MINISZTERELNÖK jött el ebbe a faluba.
the prime.minister came-3sg pv this-into the village-into
'It was the prime minister who came to this village.'

Contrastive focus in this case operates on a contextually determined set of people about whom the sentence could be true but is not: a set made up of several other persons next to the prime minister, like the president, the pope, etc. Contrastive focus identifies the prime minister among these as the only individual about whom the proposition is true. The same set-formation mechanism takes place in exclamatives. (74) forms a contrastive set that ranges over people of various significance (a janitor, a mayor, a minister or the prime minister), just like (73) does:

(74) (Hogy) ki jött el ebbe a faluba!
comp who came-3sg pv this-into the village-into
'What a person came to this village!'

[16] Contrastive focus is referred to as "identificational focus" in É. Kiss (1998).

The only difference between the lexical focus example in (73) and the exclamative one in (74) is that the set in the latter is placed on an evaluative **scale**. Scalar readings are inherent to exclamatives as was pointed out in the previous section. While this scalar reading is missing in (73), it can be brought in with the use of so-called focus sensitive adverbials like the scalar *only* or *even*. These also associate with scales (Rooth 1992) that are similar to those found in exclamatives:

(75) Csak a helyi polgármester jelent meg a faluban.
only the local mayor showed-3sg up the village-in
'Only the local mayor showed up in the village (and not a minister or the prime minister).'

A sentence like (75) establishes a scale on which the local mayor is associated with a low value of importance, as opposed to other individuals with a higher value of importance. In other words, contrast in the case of exclamatives falls on a degree property: it singles out an extreme degree, opposing it to other, less extreme degrees.

On the basis of these parallels I put forward the claim that it is the set invoking nature of exclamatives that explains their focal placement. Exclamative constituents are focused phrases, focus falling on a value of an evaluative scale.[17] The scalar nature of exclamatives can be implemented by the presence of an exclamative operator (Op_{EX}) that is present on every E-phrase as illustrated in (76). The workings of this operator is similar to that of *only* in that it establishes an evaluative scale and forces focus syntax onto the E-constituent.

(76) [... [$_{FocP}$ [$_{E\text{-}phrase}$ Op_{EX} {*ki / mi / hol / mikor / hogyan*}] V^0 [$_{AspP}$ pv ...]]]

The presence of such an evalutative scalar operator in exclamatives can be motivated indirectly by observations about the lexical specification of E-phrases. Since the exclamative operator is a scalar operator, it is expected, that other, non-scalar operators are ruled out in E-phrases. This expectation is borne out. Non-scalar focal particles, like *pontosan* 'exactly, precisely' are ruled out in E-phrases:

[17] That focus underlies exclamatives is not entirely new, although it has never been spelled out extensively in the literature to my knowledge. For proposals that touch on the focal nature of exclamatives, see Gutiérrez-Rexach (1999) and Nelson (1997) on Spanish and English exclamatives respectively.

(77) (Hogy) (*pontosan) mikor jött meg a vonat!
comp exactly when arrived-3sg pv the train
'(I am surprised at) the exact time when the train arrived.'

The syntactic placement of E-phrases occurring with inversion is thereby explained. The focus placement follows from their interpretation as exclamatives, invoking a set of alternatives that range over degree expressions. The scalar set invoking reading of exclamatives characterizes every exclamative phrase, including those that can also occur without inversion in Hungarian, like the amount phrases like *hány/mennyi* 'how much/many'. Since their exclamative interpretation does not differ in any way from other E-phrases in terms of scalar set-invoking behaviour, I take these elements to be associated with an evaluative scalar Op_{EX} operator as well. Their placement in the *many*P position (cf. section **3.3.2**), indicated in (78) does not contradict the claim that E-phrases are focused, as *many*P can contain semantically focused expressions and is evaluative by definition:[18]

(78) [... [*many*P [E-phrase Op_{EX} {*hány/mennyi* / *ki mindenki* / %*kik*} [AspP pv-V [...]]]]

Note that the focus analysis of E-phrases gets support in other domains of grammar as well as other syntactic constructions. The effects of focusing can be observed in (i) phonological behaviour; (ii) the existence of non-sentential exclamatives and (iii) relativized exclamatives. The remainder of this section elaborates on these phenomena.

Turning first to the syntax–phonology interface, the focus analysis of E-phrases is clearly supported by phonological considerations. Exclamations always contain an emphatic phrase. The emphatic phrase always corresponds to the constituent that denotes the surprising thing or property (the E-phrase), which always receives the main accent of the sentence. The focus-analysis of exclamatives thus in turn explains why exclamatives can only occur in positions which can carry the main accent of the clause, as was stated in the generalization in (36), repeated here:

(79) The position of adverbial E-phrases in exclamatives is the same as the position they occupy in indicatives when they carry the main stress of the sentence.

[18] See arguments in Brody (1990) and Surányi (2002) to the effect that inhabitants of DistP positions can have focus semantics.

This requirement follows from the fact that only positions associated with main stress can host focused constituents.

Turning now to the syntactic domain, two important exclamative constructions support the focal analysis of exclamatives. Both involve constructions that arguably involve ellipsis: non-sentential exclamatives, i.e., exclamative utterances that do not involve a full sentence. Hungarian has two kinds of non-sentential exclamatives. One involves ordinary E-phrases, like the exclamative in (72) above, repeated here:

(80) Mi-*(csoda) egy fickó!
what-wonder a guy
'What a guy!'

Like (80), and its English version, non-sentential exclamatives are usually nominative/adjectival predicative phrases. Although such utterances are widespread in languages, their non-sentential nature has not received attention in the literature. Following recent findings about so-called (non-sentential) **fragment** phrases in general (Merchant 2004), it looks likely that non-sentential exclamatives are elliptical phrases in which only one constituent survives and all other parts of the utterance are deleted. Although further research is needed to establish whether this is really so,[19] such a deletion analysis would provide a strong piece of evidence in favour of the focused nature of E-phrases: as Tancredi (1992) and Merchant (2001) have shown, ellipsis can only operate in sentences where the non-elided remnant constituent is focused. If exclamative phrases like (80) can survive ellipsis, it is clearly because they are focused.

A similar kind of reasoning can also be applied to the other kind of non-sentential exclamatives, relativized exclamatives, like (81), which was introduced in example (3) above:

(81) (a) Amennyit te egyszerre megeszel!
rel-how.many-acc you once pv-eat-2sg
'The amount you eat in one sitting!'

(b) Amiket te megeszel!
rel-what-pl-acc you pv-eat-2sg
'The things you eat!'

[19] Initial investations in English point to this direction: Ono (2005) shows that exclamative constituents in English can occur as remnants left behind in sluicing, for example.

As is clear from the morphology of the *wh*-expression, these sentences are relative clauses. Since relative clauses are dependent, subordinated constituents, these utterances, too, are most presumably non-sentential constituents just like the elliptical type in (80) involving a non-relative E-phrase. The deletion process that has taken place in relativized exclamatives arguably deletes the predicate whose subject is a lexical DP containing the relative clause:[20]

(82) (a) ~~Elképesztő a mennyiség~~, amennyit te egyszerre megeszel.
astonishing the amount rel-how.many-acc you once pv-eat-2sg
'The amount you eat in one sitting is astonishing.'

(b) ~~Elképesztőek az ételek~~, amiket te megeszel.
astonishing-pl the food-pl rel-what-pl-acc you pv-eat-2sg
'The things you eat are astonishing.'

If such a deletion analysis is on the right track,[21] this provides evidence that the relative clause (the E-phrase), is focused. Interestingly, one can find other indication that the relative clause as a whole is focused in these sentences, and that it has a particular, scalar focus interpretation that is different from ordinary contrastive focus. The evidence comes from the distribution of lexical focused phrases within the relative clause. The argumentation takes several steps, which are sketched in the following.

As the next examples show, the presence of lexical focus inside the exclamative phrase gives a sharply ungrammatical result:

[20] To some degree, the head of the relative clauses can also survive the deletion process, giving rise to headed relative exclamatives, although the result is dispreferred to the free relatives in (81). Nominal relatives are the best here (ib):

(i) (a) ???A mennyiség amennyit te egyszerre megeszel!
the amount rel-how.many-acc you once pv-eat-2sg
'The amount you eat in one sitting!'

(b) ?Az ételek, amiket te megeszel!
the food-pl rel-what-pl-acc you pv-eat-2sg
'The things you eat!'

[21] While the deletion account seems plausible for Hungarian, languages might differ in whether they derive these sentences with ellipsis or not. See some arguments against a deletion account of English *The things you eat!* type of exclamatives in Portner–Zanuttini (2005).

(83) (a) *Amennyit TE eszel meg egyszerre!
rel-how.many-acc you eat-2sg pv once
'The amount YOU eat in one sitting (as opposed to someone else)!'

(b) *Amiket TE eszel meg!
rel-how.many-acc you eat-2sg pv
'The things YOU eat (as opposed to someone else)!'

This behaviour is particular, as relative clauses in general can easily contain lexical focus expressions:

(84) [Amiket TE eszel majd meg], AZOKAT tettem a hűtőbe.
rel-how.many-acc you eat-2sg later pv those-pl placed-1sg the fridge-into
'The things YOU will eat (as opposed to someone else) are the ones I have put into the fridge.'

The obligatory absence of a focus phrase inside the relative in (83) follows from the fact that the relative clause as a whole is a focus expression, which, as was argued above in this section, is associated with a scalar operator Op_{EX}. Due to this operator, the relative clause assumes a scalar focus reading, just like any other E-phrase:

(85) [$_{\text{E-phrase}}$ Op_{EX} [amennyit te egyszerre megeszel]]!

As a result, the exclamative can only have this kind of focus reading, and is not compatible with an ordinary contrastive focus reading at the same time. The problem that results when the relative clause contains a lexical focus is precisely the latter: the relative receives two conflicting types of focus interpretation. This is due to a phenomenon called **focus percolation** that characterizes Hungarian relative clauses in general. In short, focus percolation means that the interpretation and syntactic distribution of Hungarian relative clauses is sensitive to whether they **contain** a focused constituent or not. If a free relative clause contains a focus element, the relative clause as a whole strongly prefers to be focused as well. This is illustrated in (86):

(86) (a) ???*János megette [amit MARI készített el].
János pv-ate rel-what-acc Mari prepared pv
'János ate what MARI prepared.'

(b) [Amit MARI készített el] ette meg János.
rel-what-acc Mari prepared pv ate pv János
'It was what MARI prepared that János ate.'

As can be seen in these examples, the presence of focus on *Mari* within the relative clause forces focus interpretation and focus distribution on the whole relative clause as well. If the relative is not in SpecFocP (triggering inversion), the sentence becomes considerably degraded, as (86a) shows.[22] It is this focus percolation phenomenon that rules out lexical focus inside exclamative relatives. An internal focus would force contrastive focal interpretation onto the whole relative sentence, and as such it would come into conflict with the scalar focus interpretation that is a necessary ingredient of the exclamative phrase:

(87) *[$_{(+\text{ focus})}$ [E-phrase Op_{EX} [amennyit $\text{TE}_{(+\text{ focus})}$ eszel meg]]]!

Since Op_{EX} and (+ focus) are linked to a distinct contrastive focus meaning, they are in complementary distribution. This conclusion in turn provides evidence for the obligatorily focused nature of E-phrases. If the exclamative relative clause was not marked for an exclamative scalar focus reading, a lexical focus phrase could happily occur inside it, similarly to the grammatical instances of relative clauses with lexical foci where this is possible (84).

Summarizing the claims, this section spelled out a theory of exclamatives in which their syntactic focus behaviour was derived from their inherent semantic properties of being scalar and set-invoking. It was argued that similarly to focused phrases with scalar operators like *only*, Hungarian exclamative phrases contain a covert exclamative operator that gives them scalar semantics. Due to this operator they have the distribution of focused elements: they have to occur in a verbal position which is either FocP or *many*P. Evidence for the focused nature of the E-phrases was also provided from the realm of non-sentential, arguably elliptical, exclamative constructions.

[22] Next to focusing the relative clause itself (as in (86b)), there are other ways of licensing an internal focus in relatives. One involves focusing the head of the relative, if there is one (ia). Another one involves focusing some other phrase in the matrix clause (ib) (István Kenesei, p.c.):

(i) (a) János AZT ette meg [amit MARI készített el].
János that-acc ate pv rel-what-acc Mari prepared pv
'It was what MARI prepared that János ate.'

(b) JÁNOS ette meg [amit MARI készített el].
János ate pv rel-what-acc Mari prepared pv
'It was János who ate what MARI prepared.'

5. Summary of findings and the relevance of the Hungarian data for the theory of exclamatives

The aim of this last section is to summarize the results of the paper and highlight their relevance for the study of exclamative syntax in general. The latter is very important as syntactic research on exclamatives has up to now been rather small-scale, based on data from few languages only. The Hungarian data have received no theoretical attention yet in cross-linguistic research, although they are clearly relevant for the study of exclamativity in general, as they show syntactic behaviour that to my knowledge has not been attested in other languages.

5.1. Summary of findings

The present paper was devoted to the study of exclamative sentences, concentrating primarily on word order properties of the most frequent types of exclamative constructions. Hungarian E-phrases can be formed both with or without *wh*-words. Non- *wh*-E-phrases contain an intensifying element *de* 'how' and distribute just like *wh*-E-phrases.

The syntactic distribution of E-phrases in Hungarian is a complex matter that was carefully described in section **2**. E-phrases were found to fall into two major types depending on what position they occupy. The larger class can only occur in the Spec,FocP position, triggering inversion. These are phrases which do not inherently refer to an amount. A smaller class of phrases can also occur in a different position, which was argued to be a quantificational position, Spec, *many*P, the position of the evaluative amount expression *sok* 'many' in indicative clauses, which is not associated with inversion. This class of E-phrases comprises amount phrases only. The two positions, FocP and *many*P, which exclamative phrases can occupy were distinguished using syntactic tests and co-occurrence restrictions. The distribution of E-phrases can thus be summarized in (88):

(88) (a) [... [*many*P {*amount*-phrases} [AspP pv-V [...]]]]
(b) [... [FocP {*amount* and *non-amount*-phrases} V^0 [AspP pv ...]]]

Alongside these cases, there exists yet another small class with intensity expressions that always occur without inversion, due to a (sometimes covert) adverbial of intensity that can never be focused.

Similar to these intensity adverbs, E-phrases with other adverbial elements (as well as quantificational ones) have also been studied. It was shown that quantificational and adverbial phrases that always show a peculiar, stress-sensitive behaviour in the preverbal domain retain the same behaviour in exclamatives as well. These phrases were found to be placed in the same position in both indicative and exclamative clauses.

These findings lead to the conclusion that exclamative clause type is not associated with a singular position in which E-phrases have to appear. In this they starkly differ from interrogative clauses, which always host interrogative phrases in the same position, Spec,FocP.

The reason for not being associated with one syntactic position is that exclamatives are associated with a type of focal semantics that can be hosted in more than one position: evaluative scalar focus. Due to the fact that evaluative scalar focus phrases can be hosted by more than one syntactic slot, exclamatives can be placed in more than one slot as well: in FocP, *many*P or preverbal adverbial positions. The choice between these positions is entirely lexical. *Many*P only accepts amount expressions, while FocP is compatible with both amount phrases and other types. Adverbial positions in exclamatives correspond to adverbial positions in indicatives and are selective for features like +grade, +contrast, etc.

These results prove relevant for various domains of syntactic investigations, the most important Hungarian-specific one of these being the study of clause types, focus types and the layout of the left periphery. In addition, the findings also have important ramifications for the cross-linguistic theory of exclamatives, as the next section will show.

5.2. Relevance of findings for the study of exclamativity

The Hungarian facts discussed in this paper have important consequences for the syntatic study of exclamatives. This is the more important to point out, as exclamatives have not yet been extensively studied. The handful of articles on this topic usually concentrate on one or the other exclamative construction in a given language (Postma 1996; Bennis et al. 1998; Bennis 1998 on Dutch; D'Avis 2002 on German; Grimshaw 1979; Nelson 1997; Pesetsky–Torrego 2001; Portner–Zanuttini 2003; Fujii–Ono 2005 on English; Ono 2002 on Japanese; Portner–Zanuttini 2003; Munaro 2003 on Italian (Paduan and Bellunese respectively); Espinal 1997; Gutiérrez-Rexach 1999; Villalba 2003 on Spanish). Theoretical studies building on a cross-linguistic approach are missing, except for pioneering

work in Portner–Zanuttini (2003), which is the only study providing a general theory of exclamatives (based on an English–Italian comparison). It is the claims of this latter paper that will be reviewed in the light of the Hungarian data discussed above.

The part of Portner–Zanuttini (2003) (P & Z for short) for which the Hungarian facts are relevant concern the definitions provided for exclamatives. Based on the crucial meaning components of exclamations, P & Z design tests with which exclamatives can be distinguished from other clause types and can be defined accurately in the syntax.

One of the important syntactic properties that characterize exclamatives in their view is their operator-variable structure. The operator-variable structure gives rise to the fact that exclamatives denote alternative propositions, as was shown in section **4.1** above. Portner and Zanuttini moreover define the operator-variable structure found in exclamatives as a necessary-*wh* operator-variable structure, and build their semantics explicitly on *wh*-quantification in them:

(89) Exclamatives involve a *wh* operator-variable structure. (= (1a) in P & Z)

Hungarian exclamatives provide evidence that *wh*-syntax is not a necessary ingredient of exclamatives. In addition to *wh*-exclamatives, *de*-exclamatives also have the exact same distribution and meaning as *wh*-exclamatives. An example from above (21b) demonstrates this again:

(90) {De /milyen} sok könyvet {megvettél / vettél meg}! (= (21b))
de how many book-acc pv-bought-2sg bought-2sg pv
'You bought so many books!'

It is clear therefore that the definition in (89) is too narrow to cover all data. The Hungarian facts show that languages can use available syntactic means of focus in the expression of exclamatives. The minimal modification to be made in (89) is the addition that the operator-variable structure can also be an operator-variable structure of the focus kind.

This modification is by no means a substantial amendment, as it is known from the literature that contrastive focus also sets up an operator-variable structure, just like *wh*-movement (É. Kiss 1998). As was spelled out above, contrastive focus also requires the consideration of a set of alternatives (Rooth 1992), just like *wh*-operators do. Defining the operator-variable structure of exclamatives in terms of focus is also advantageous because it subsumes the earlier proposal in terms of a *wh*-variable struc-

ture (cf. 89), if one subscribes to the view that *wh*-movement is a subcase of focusing (Chomsky 1977; Rochemont 1986). Adopting the latter view, the new definition can be given in (89′):

(89′) Exclamatives involve a focus operator-variable structure.

The second characteristic of exclamatives in P & Z is the factivity of exclamatives (see also Fujii–Ono 2005 for an analysis that uses factivity for the syntactic definition of exclamatives). To use factivity as a test for exclamatives, P & Z propose to identify exclamatives as *wh*-clauses that cannot be embedded under non-factive predicates (following Grimshaw 1979):

(91) Mary knows / *thinks / *wonders how very cute she is.

While the factivity of exclamatives is beyond doubt, and the embedding test is sound, the problem that Hungarian presents with respect to this test is its inapplicability to exclamative constructions that in general cannot be embedded, namely all *de*-exclamatives in Hungarian:[23]

(92) (a) *Elképesztő, hogy de rohadt hideg van. (= (13b, c))
astonishing that *de* rotten cold is
'It's astonishing how awfully cold it is.'

(b) *Meglep, hogy de rohadt hideg van.
surprise-3sg that *de* rotten cold is
'It surprises me how awfully cold it is.'

The question is then, how to characterize *de*-exclamatives. What differentiates them from *wh*-exclamatives and what rules them out in embedded positons?

Notice that disqualifying *de*-exclamatives from exclamativehood would not do. Native speaker intuition "feels" that if something, *de*-exclamatives are even more strongly "exclamative" in the illocutionary sense of this word than *wh*-exclamatives. I propose to implement this intuition by saying that *de*-exclamatives and *wh*-exclamatives differ in their **expressivity**.

Expressivity characterizes expressive constructions, the latter being phrases that are strongly tied to the utterance situation, notably to

[23] The same problem is presented by certain Spanish exclamatives (Espinal 1997) or Japanese ones (Ono 2002), which cannot be embedded, either.

the emotional state of the speaker (Potts 2005; Potts–Roeper 2006).[24] Expressivity in exclamatives is present due to their semantic property **emotiveness** (see section **4.1** above). While some kind of emotiveness is present in every exclamative, it seems to be the case that not all exclamatives have the same emotive content or expressive nature. Certain exclamatives are always expressive, while others only optionally are.

With this assumption in place, we can explain the distribution of *de*-exclamatives in (91), if we make the further plausible assumption that **expressive exclamatives** are only licensed in root contexts—due to the fact that only root contexts have exclamative illocutionary force (defined as a speech act in the sense of Austin 1962), which I take to be licensing expressive exclamatives. As a result, expressive exclamatives can only be used in root clauses. If this argumentation is on the right track, the distinction between *de*-exclamatives and *wh*-exclamatives in Hungarian boils down to a difference in expressivity: the former are obligatorily expressive, while the latter are not.

It must be noted that the obligatory root occurrence of expressive exclamatives can also be witnessed in the case of *wh*-exclamatives. When *wh*-exclamatives are embedded under a predicate that cannot express the same utterance situation, they loose their expressivity. Consider the following examples:

(93) (a) How very tall Lisa is! (expressive utterance)

(b) John finds it amazing how very tall Lisa is. (non-expressive utterance)

While (92a) is an expressive statement about the speaker's surprise, (92b) is not expressive either about the speaker's surprise, nor about John's. This can be easily captured by saying that *wh*-exclamatives in embedded positions are not expressive.

If these conclusions are on the right track, they provide evidence that expressivity, an optional property of exclamatives, is a syntactically relevant notion that needs to be used in the characterization of exclama-

[24] Expressive adjectives for example can only be attributed to the speaker, even when they are embedded. Consider for example the expressive *damn* in (i) (Potts–Roeper *ibid.*):

(i) [Bill reporting to Sue]: "John says that your *damn* dog has bitten the neighbour's cat!"

Even though *damn* is found in the clause that reports John's statement, it indicates the speaker's (Bill's) disapproval. This shows that expressive elements are linked to the utterance situation, not their syntactic environment.

tive constructions. This enables a finer distinction between diverse types of exclamative contexts and constructions both within one language and across languages.

References

Austin, John L. 1962. How to do things with words. Oxford University Press, Oxford.

Bennis, Hans 1998. Exclamatives! In: Renée van Bezooijen – René Kager (eds): Linguistics in the Netherlands 1998, 27–40. John Benjamins, Amsterdam & Philadelphia.

Bennis, Hans – Norbert Corver – Marcel den Dikken 1998. Predication in nominal phrases. In: Journal of Comparative Germanic Linguistics 1 : 85–117.

Brody, Michael 1990. Remarks on the order of elements in the Hungarian focus field. In: István Kenesei (ed.): Approaches to Hungarian 3 (Structures and arguments), 95–121. JATE, Szeged.

Brody, Michael 1995. Focus and Checking Theory. In: István Kenesei (ed.): Approaches to Hungarian 5 (Levels and structures), 29–43. JATEPress, Szeged.

Cheng, Lai-Shen 1991. On the typology of *wh*-questions. Doctoral dissertation, MIT.

Chomsky, Noam 1977. On *wh*-movement. In: Peter W. Culicover – Thomas Wasow – Adrian Akmajian (eds): Formal syntax, 71–132. Academic Press, New York.

Chomsky, Noam 1995. The minimalist program. MIT Press, Cambridge MA.

Corver, Norbert 1990. The syntax of left branch extractions. Doctoral dissertation, Tilburg University.

D'Avis, Franz-Josef 2002. On the interpretation of *wh*-clauses in exclamative environments. In: Theoretical Linguistics 28 : 5–32.

Dikken, Marcel den – Anikó Lipták 1997. Csoda egy nyelv — Nominal internal predication in Hungarian. In: Jane Coerts – Helen de Hoop (eds): Linguistics in the Netherlands 1997, 61–72. John Benjamins, Amsterdam & Philadelphia.

É. Kiss, Katalin 1987. Configurationality in Hungarian. Akadémiai Kiadó, Budapest.

É. Kiss, Katalin 1998. Identificational focus versus information focus. In: Language 74 : 245–73.

É. Kiss, Katalin – Donka Farkas 2001. On the comparative and absolute readings of superlatives. In: Natural Language and Linguistic Theory 18 : 417–55.

Elliott, Dale E. 1974. Toward a grammar of exclamations. In: Foundations of Language 11 : 231–46.

Espinal, Teresa Maria 1997. Non-negative negation and *wh*-exclamatives. In: D. Forget – P. Hirschbuehler – F. Martineau – M.-L. Rivero (eds): Negation and polarity, 75–93. John Benjamins, Amsterdam & Philadelphia.

Fujii, Tomohiro – Hajime Ono 2005. English *wh*-exclamatives and T-to-C movement. Manuscript. University of Maryland. (http://www.ling.umd.edu/hajime/)

Grimshaw, Jane 1979. Complement selection and the lexicon. In: Linguistic Inquiry 10 : 279–326.

Gutiérrez-Rexach, Javier 1999. Spanish exclamatives and the interpretation of the left periphery. In: Johan Rooryck and Yves D'Hulst (ed.): Romance languages and linguistic theory, 167–94. John Benjamins, Amsterdam & Philadelphia.

Horvath, Julia 1981. Aspects of Hungarian syntax and the theory of grammar. Doctoral dissertation, UCLA.

Horvath, Julia 2000. Interfaces vs. the computational system in the syntax of focus. In: Hans Bennis – Martin Everaert (eds): Interface strategies, 183–206. Holland Academic Graphics, Amsterdam.

Kálmán, László 2001. Magyar leíró nyelvtan. Mondattan I [Hungarian descriptive grammar. Syntax 1] (Segédkönyvek a nyelvészet tanulmányozásához VI [Supplementary textbooks on linguistics 6]). Tinta Kiadó, Budapest.

Kenesei, István 1986. On the logic of Hungarian word order. In: Werner Abraham – Sjaak de Meij (eds): Topic, focus, and configurationality, 143–159. John Benjamins, Amsterdam.

Kiefer, Ferenc 1967. On emphasis and word order in Hungarian. Indiana University Press, Bloomington.

Koopman, Hilda – Anna Szabolcsi 2000. Verbal complexes. MIT Press, Cambridge MA.

Lipták, Anikó 2001. On the syntax of *wh*-items in Hungarian. Doctoral dissertation, LOT, Leiden.

Lipták, Anikó 2005. The left periphery of Hungarian exclamatives. In: Laura Bruge – Giuliana Giusti – Nicola Munaro – Walter Schweikert – Giuseppina Turano (eds): Proceedings of 30th "Incontro di Grammatica Generativa", 161–83. Cafoscarina, Venezia.

Merchant, Jason 2001. The syntax of silence. Oxford University Press, Oxford.

Merchant, Jason 2004. Fragments and ellipsis. In: Linguistics and Philosophy 27: 661–738.

Michaelis, Laura A. – Knud Lambrecht 1996. The exclamative sentence type in English. In: Adele Goldberg (ed.): Conceptual structure, discourse and language, 375–89. CSLI Publications, Stanford, CA.

Munaro, Nicola 2003. On some differences between exclamative and interrogative *wh*-phrases in Bellunese: Further evidence for a Split-CP Hypothesis. In: Christina Tortora (ed.): The syntax of Italian dialects, 137–51. Oxford University Press, Oxford.

Nelson, Nicole 1997. The structure of exclamatives in English. Manuscript. Ruthers University.

Obenauer, Hans 1976. Etudes de syntaxe interrogative du français. Niemeyer, Tübingen.

Olsvay, Csaba 2000. Formális jegyek egyeztetése a magyar nemsemleges mondatokban [The checking of agreement features in Hungarian non-neutral sentences]. In: László Büky – Márta Maleczki (eds): A mai magyar nyelv leírásának újabb módszerei IV [Recent methods in the description of contemporary Hungarian 4], 119–52. JATEPress, Szeged.

Ono, Hajime 2002. Exclamatory sentences in Japanese: A preliminary study. In: Yukio Otsu (ed.): Proceedings of the Third Tokyo Conference on Psycholinguistics, 305–26. Hituzi Syobo Publications, Tokyo.

Ono, Hajime 2005. Sluicing in exclamatives. Paper presented at the ECO5 Syntax Workshop, Harvard University. Available at: http://www.ling.umd.edu/hajime/.

Pesetsky, David – Esther Torrego 2001. T-to-C Movement: Causes and Consequences. In: Michael Kenstowicz (ed.): Ken Hale: A life in language, 355–426. MIT Press, Cambridge MA.

Portner, Paul – Raffaela Zanuttini 2003. Exclamative clauses: At the syntax-semantics interface. In: Language 79 : 39–81.

Portner, Paul – Raffaela Zanuttini 2005. The semantics of nominal exclamatives. In: Robert J. Stainton – Reinaldo Elugardo (eds): Ellipsis and non-sentential speech, 57–67. Kluwer, Dordrecht.

Postma, Gertjan 1996. The nature of quantification of high degree: 'very', 'many' and the exclamative. In: Crit Cremers – Marcel den Dikken (eds): Linguistics in the Netherlands, 207–20. John Benjamins, Amsterdam & Philadelphia.

Potts, Christopher 2005. The logic of conventional implicatures. Oxford University Press, Oxford.

Potts, Christopher – Tom Roeper 2006. The narrowing acquisition path: From expressive small clauses to declaratives. In: Ljiljana Progovac – Kate Paesani – Eugenia Casielles – Ellen Barton (eds): The syntax of non-sententials: Multi-disciplinary perspectives, 183–201. John Benjamins, Amsterdam & Philadelphia.

Rochemont, Michael 1986. Focus in generative grammar. John Benjamins, Amsterdam.

Rooth, Mats 1992. A theory of focus interpretation. In: Natural Language Semantics 1 : 75–116.

Sadock, Jerrold M. – Arnold Zwicky 1985. Speech act distinctions in syntax. In: Timothy Shopen (ed.): Language typology and syntactic description, 155- 196. Cambridge University Press, Cambridge.

Surányi, Balázs 2002. Multiple operator movements in Hungarian. Doctoral dissertation, LOT, Utrecht.

Szabolcsi, Anna 1997. Strategies for scope taking. In: Anna Szabolcsi (ed.): Ways of scope taking (SLAP 65), 109–155. Kluwer, Dordrecht.

Tancredi, Christopher 1992. Deletion, de-accenting, and presupposition. Doctoral dissertation, MIT.

Villalba, Xavier 2003. An exceptional exclamative sentence type in Romance. In: Lingua 118 : 713–45.

Acta Linguistica Hungarica, Vol. 53 (4), pp. 393–432 (2006)
DOI: 10.1556/ALing.53.2006.4.2

SCRAMBLING IN HUNGARIAN*

BALÁZS SURÁNYI

Research Institute for Linguistics
Hungarian Academy of Sciences
Benczúr utca 33.
H-1068 Budapest
Hungary
suranyi@nytud.hu

Abstract: This article revisits the (non)configurationality debate of the 80s and early 90s concerning Hungarian, a "free word order" language, which was shown during that period to be characterized by an articulate and, crucially, hierarchical preverbal domain, with A-bar positions dedicated to discourse functions such as topic and focus. What this debate did not conclusively settle, however, is the question whether or not the structure of A-positions in Hungarian is also configurational. The most prevalent, and indeed empirically most well-argued and elaborated analysis that has emerged is that of É. Kiss's (1987a; 1987b; 1991; 1994b; 2002; 2003), according to which the answer

* It has taken what appears now an unthinkably long gestation period for the ideas presented here to materialize in the form of a paper. The first person to thank here is Elsi Kaiser, with whom I first discussed some details of a scrambling approach to the Hungarian postverbal domain at ConSOLE 7 in Bergen in 1998. I am especially indebted to Máté Bánki, not only for his interest in this topic as a student, which made me resume the work on it which I left off sometime during my graduate years, but also for kindly performing an extensive pilot test during the Fall semester of 2004 to collect judgments on some of the sentence types I draw on. I also thank my own linguist and non-linguist informants for their time and patience with my questionnaires. I would like to thank Katalin É. Kiss and members of her project group at the Research Institute for Linguistics in Budapest, where the analysis proposed here was first presented in April 2005, for helpful questions and comments. Thanks are also due to the audience at NELS 36 at UMass, Amherst, where part of this material was also presented. All errors, of course, are my own. The work reported here has been supported by grant No. D-048454 of the Hungarian Scientific Research Fund, which is hereby gratefully acknowledged.

1216-8076/$ 20.00

is negative: arguments are base-generated in the verb phrase in a free order in a flat structure. The present paper challenges this view by demonstrating systematically that the arguments put forward to back it up are inconclusive, and in fact it fails descriptively as well. The alternative proposed here is based on a hierarchical verb phrase (vacated by the raised verb) and a Japanese-type local scrambling movement that operates in the post-verbal domain of the clause. The scrambling movement analysis, besides being theoretically more desirable than the nonconfigurational verb phrase approach, makes available a superior descriptive coverage by accounting for a varied set of structural symmetries and asymmetries holding between subject and object. Modulo scrambling, Hungarian is configurational all the way down.

Keywords: scrambling, (non)configurationality, Weak Crossover, binding, Hungarian

1. Introduction

Hungarian is well-known for its overt movements to a richly articulated preverbal left-periphery ("discourse-configurationality"), where syntactic hierarchy and scope interpretation are isomorphic (e.g., É. Kiss 1987a; 1991; 1995; Kenesei 1986). By contrast, its postverbal domain, where constituent order exhibits a radical freedom, has received much less attention. The only empirically argued and elaborated analysis that has emerged from the discussion in the 80s/early 90s of the issue whether Hungarian is configurational with respect to its argument positions is that of É. Kiss's (1987a; 1987b; 1991; 1994b; 2002; 2003). According to this view, the Hungarian nuclear clause (S in (1987a; 1987b) and later, VP) is non-configurational (cf. also Kenesei 1984), in particular, arguments (and adjuncts) are generated in a free order in a flat structure, as schematized in (1) (É. Kiss 1994b; 2002; 2003).[1]

(1) (a) $[_{VP}\ [_{V'}\ V\ DP_{subj}\ DP_{obj}]]$
(b) $[_{VP}\ [_{V'}\ V\ DP_{obj}\ DP_{subj}]]$

The flat VP analysis is not without appeal due to its descriptive merits, which is probably the reason why it has become the most widely

[1] In É. Kiss (1994a), the VP is flat and there are no inflectional projections like AgrPs or TP in the clause. Her (2002) survey of Hungarian syntax does adopt inflectional projections for the treatment of inflectional morphology, but these projections are assumed to play no role in the syntax of arguments.

A terminological caveat is also in order: permutation in the flat VP has also been referred to as 'scrambling' in the literature on Hungarian, but this sense of the term should be kept apart from the claim made in the present paper.

accepted view in the literature on Hungarian.[2] In this paper I argue for an alternative approach in terms of a hierarchical basic VP structure, coupled with local scrambling movements. It is demonstrated that the central subject–object symmetries that have served as the empirical base for the flat VP analysis can be accommodated equally well under the hierarchical VP plus scrambling approach, while on the other hand, a number of subject–object asymmetries (to be pointed out) strongly favor the latter view.

The structure of the paper is as follows. In section **2** I enumerate the empirical arguments that have been presented in favor of a non-hierarchical VP structure, which involve a lack of subject–object (S/O) asymmetries.[3] In section **3** I demonstrate that the arguments reviewed in section **2** in favor of a flat VP are inconclusive: some of the arguments are ill-founded, and some others lose force once a scrambling movement analysis based on a hierarchical VP (or rather, *v*P) is shown to derive the observed patterns equally well. Section **4** presents a host of asymmetry facts that are problematic under a flat VP analysis, but fall out on a scrambling approach modulo the hierarchical *v*P that this account adopts. Section **5** examines the basic properties of the postverbal object–subject reordering in Hungarian, and demonstrates that this reordering is akin in particular to (local) scrambling of the Japanese-type. Section **6** concludes the paper, and spells out the significance of its results.

[2] This is not to say that the configurational view has lacked proponents (e.g., Horváth 1986; Marácz 1989, and Speas 1990; cf. also Kenesei 1989). Nevertheless, the proposed implementations of a configurational approach were partly incomplete and partly descriptively inadequate, and/or relied on analytic devices that are no longer available (or, in some cases, even formulable) in the current restrictive framework. In a recent manuscript I received while working on this material, Katalin É. Kiss (2006) offers a phase-based analysis of the Hungarian verb phrase, which adopts the view that the verb phrase is hierarchical at the initial point of the derivation, but maintains a flat structure for the verb phrase at some later point in the derivation. I must leave the comparison of this recent hybrid approach with the one put forward in the present paper for another occasion.

[3] To keep a reasonable depth of subject matter, I will limit the discussion to the base position of the subject and the direct object; the placement of postverbal internal arguments and adjuncts cannot be addressed within the confines of this paper. Nevertheless, the scrambling operation envisaged to apply in Hungarian displaces not only direct objects, but also other internal arguments, including not only DPs but PPs as well.

2. The coverage of the flat VP analysis

The flat VP approach is designed to capture two central properties of the syntax of the Hungarian clause: first, constituent order to the right of the verb exhibits a degree of freedom unattested in "fixed word order" languages like English; and second, Hungarian is assumed to lack most subject/object asymmetries characteristic of languages where the subject is base-generated in a position higher than the object. This is not to say that Hungarian lacks subject/object asymmetries altogether. Within a flat VP approach, observable asymmetries of that kind must be treated as non-structural in nature, as we will see shortly. In this section I review the major arguments cited to back up the non-configurational view of the Hungarian verbal phrase (É. Kiss 1987a;b; 1994b; 2002; 2003).

2.1. Weak Crossover

(i) Weak Crossover (WCO) effects typically obtain when an element X is A-bar moved across an expression Y, where Y properly contains a variable bound by X. In languages like English, where the A-position of the subject is higher than the A-position of the object (the former c-commands the latter), *wh*-movement of the object across the subject gives rise to a marked degradation in acceptability, as in (2a) below, while no such deterioration is observable in the reverse configuration, as in (2b). Reinhart (1983) proposes to capture WCO in terms of the configuration licensing bound variable pronouns: bound variables must be A-bound (bound from an A-position) by their antecedent. Koopman and Spotiche's (1982) alternative view based on their Bijection Principle essentially dictates that pronouns must not be locally A-bar bound. According to Lasnik and Stowell's (1991) formulation, if a pronoun *pron* and a trace t left behind by an A-bar movement are both bound by a quantifier (like the displaced *wh*-expression), then t must c-command *pron*. It follows from (2) on any one of these generalizations that the A-position of the object does not c-command the A-position of the subject (and the pronoun inside it), while the A-position of the subject does c-command the A-position of the object (and hence also the pronoun inside it).

(2) (a) ?*[Who$_i$ does [[his$_i$ mother] like t_i]]?
(b) [Who$_i$ [t_i likes [his$_i$ mother]]]?

Hungarian, by contrast, is a language that does not display a WCO effect in analogous constructions (see (3a)), which has received an explanation under the flat VP analysis as follows. If the VP is flat, the position (marked by t_i) from which the object is *wh*-moved across the subject is sister to the position of the subject. No WCO effect obtains, because the object is moved from a position where it c-commands the (co-indexed pronominal variable inside the) subject. The WCO effect is absent also when the subject undergoes *wh*-movement, as is the case in English. On the flat VP analysis, this is expected as the A-position of the object (and hence the pronoun inside it) is c-commanded by the A-position of the subject.

(3) (a) $[\text{Kit}]_{\text{ACC}i}$ hívott fel [az pro_i anyja]$_{\text{NOM}}$ t_i?
who-acc$_i$ called-3sg up the (his$_i$) mother-poss.3sg-nom
'?*Who$_i$ did his$_i$ mother call up?'

(b) $[\text{Ki}]_{\text{NOM}i}$ hívta fel t_i [az pro_i anyját]$_{\text{ACC}}$?
who-nom$_i$ called-3sg up the (his$_i$) mother-poss.3sg-acc
'Who$_i$ called up his$_i$ mother?'

WCO effects are not wholly absent from Hungarian: they are attested in long *wh*-movement, as illustrated by the contrasted sentences below.[4]

[4] This fact is exemplified by Puskás (2000, 293), however, her example is ungrammatical independently of the WCO configuration; the one in (4a) is out exclusively due to WCO (the degradation is only aggravated by whatever factor determines long *wh*-movement to be felt marked by many speakers of Hungarian, compare (4b)). Puskás's example is (i). (i), however, is independently rendered ungrammatical by the choice of affixation on the embedded verb, which in (i) agrees with a definite object (*wh*-pronouns are known to trigger indefinite object agreement conjugation on the selecting verb). But even granting the correct (indefinite) agreement form of the embedded verb, the example is out (even when it involves a matrix subject that does not contain a bound pronominal, cf. (ii), where *pro* in indexical), because the matrix verb form is also incompatible in (i) with an object *wh*-pronoun long-moved into the matrix clause, which routinely triggers indefinite object agreement on the matrix verb.

(i) *Kit$_i$ mondta az pro_i anyja, hogy a fiúk látták t_i?
who-acc said.3sg.defobj the (his) mother-nom that the boys-nom saw.3pl.defobj
'Who did his mother say the boys had seen?'

(ii) *Kit$_i$ mondta az *pro* anyád, hogy a fiúk láttak t_i?
who-acc said.3sg.defobj the **your** mother-nom that the boys-nom saw.3pl.**indefobj**

Note that the acceptability of long *wh*-movement is known to exhibit a certain degree of variation among speakers: the spectrum goes from speakers who find them perfectly acceptable and also use them in their speech to those that flatly

In the (a) example, long *wh*-movement proceeds across the dative DP in the matrix clause, which embeds a silent pronoun co-referring with the moved *wh*-element. In (b), in contrast, the deictic second person covert pronoun *pro* does not interfere.

(4) (a) *Kivel$_i$ mondtad az *pro*$_i$ anyjának, hogy
who-with say-past-2sg.def the (his) mother-poss.3sg-dat that
kikezdtek a fiúk t_i?
flirted-3pl the boys-nom
'*Who$_i$ did you tell his$_i$ mother that the boys had flirted with?'

(b) ?Kivel$_i$ mondtad az *pro* anyádnak, hogy
who-with say-past-2sg.def the (your) mother-poss.2sg-dat that
kikezdtek a fiúk t_i?
flirted-3pl the boys-nom
'Who did you tell your mother that the boys had flirted with?'

This observation is important to make, as it shows that Hungarian has no internal property which would preempt WCO effects in general; nevertheless, WCO is unattested in short *wh*-movement of objects.[5]

---

reject them. Of course, the contrast above exists only for speakers who accept long *wh*-movement constructions to begin with.

[5] Brody (1995) argues that when undergoing *wh*-movement, objects touch down in a Case-checking specifier, [Spec,AgrOP], an A-position from which the *wh*-object c-commands and A-binds the pronoun within the VP-internal subject. This is claimed to be the reason why WCO is unattested with short *wh*-movement in Hungarian. It is irrelevant that this Case-related position is identified as the *v*P-edge in the more recent AgrP-less clause structure model: what is crucial is that it has the property of being above the base position of the subject. Precisely this latter property is argued against and is rejected by the Johnson – Koizumi – Lasnik approach to object Case checking, also embraced by Chomsky in his most recent work, according to which the Case position of the object is higher than its own base position, but lower than the base position of the subject. Independently of this issue, however, there are a number of empirical problems with the suggestion. For one thing, the same WCO-obviating derivation is expected to be available with long *wh*-movement too, since long-moved *wh*-objects trigger (indefinite) object agreement on the matrix verb, which indicates that the moved *wh*-object passes through matrix AgrOP. Then, the WCO-effect is predicted to be obviated with long movement of *wh*-objects, which is contrary to fact, cf. (i) (vs. (ii)). Another inadequacy of Brody's (*ibid.*) Case-position based proposal is that it fails to extend to non-object internal arguments, which behave on a par with objects with regard to WCO, but which do not bear structural Case, and consequently are not related to a Case-checking position above the subject. Finally, as also pointed out by É. Kiss (2002), it is unclear

2.2. Superiority

(ii) Superiority effects in single *wh*-fronting languages like English are exemplified by (5). In this language type it is the higher *wh*-item that must be attracted to the left periphery, i.e., to CP. The effect of Superiority in a multiple fronting language is illustrated from Bulgarian in (6): the original c-command relations between the *wh*-elements must be preserved after multiple *wh*-fronting (see Bošković 2002; Richards 1997).

(5) (a) Who saw what?
(b) *What did who see?

(6) (a) Koj kogo vižda?
who-nom who-acc sees
'Who sees whom?'
(b) *Kogo koj vižda?

Hungarian shows no sensitivity to Superiority in multiple *wh*-fronting:

(7) (a) Ki mit vett?
who-nom what-acc bought
'Who bought what?'
(b) Mit ki vett?
what-acc who-nom bought
'What was bought by whom?'

If neither argument is generated higher than the other, the lack of a Superiority effect in multiple *wh*-fronting of arguments of the same verb is expected.

---

why the same derivation (with an identical WCO-obviation effect) does not occur in English.

(i) *Kit$_i$ mondott az *pro*$_i$ anyja, hogy megláttak t_i?
who-acc said.3sg.indefobj the (his) mother-nom that pv-saw.3pl.indefobj
'Who did his mother say that they had seen?'

(ii) ?Kit$_i$ mondott az *pro* anyád, hogy megláttak t_i?
who-acc said.3sg.indefobj the (your) mother-nom that pv-saw.3pl.indefobj
'Who did your mother say that they had seen?'

2.3. Idioms and compositional theta-role assignment

(iii) A particularly interesting variety of evidence offered in favor of the flat VP analysis comes from idiom chunks. É. Kiss points out that as opposed to configurational languages of the English type, besides V+O idioms, Hungarian also has idioms composed of V+S, exemplified below.

(8) (a) Jánosra rájár a rúd (pv = preverb/verbal particle)
J-onto pv-goes the stick-nom
'John is having trouble.'

(b) Jánost eszi a fene
J-acc eats the plague-nom
'John is extremely worried/envious.'

In other words, there is no subject/object asymmetry in the domain of idiom chunks either, as predicted by a non-hierarchical VP structure.

Similar in vein is the argument from indirect theta role assignment. English has numerous examples where the theta role of the subject is determined not simply by the verb, but by the choice of verb and object together, as in (9). Now given that English is taken to be characterized by a lack the opposite scenario, where the choice of the subject would determine the semantic role of the object, this has been taken to constitute evidence that the subject is external to a constituent containing the verb and the object (VP, prior to the VP-Internal Subject Hypothesis (VISH), V′ or big VP after the VISH became generally accepted) (cf. Marantz 1984).

(9) (a) John broke a vase
(b) John broke an arm

É. Kiss (1987b, 244), citing Komlósy (1983), points out examples from Hungarian, where it is the choice of the subject that determines the semantic role of the object:

(10) (a) Eszi Jánost az oroszlán
eats J.-acc the lion-nom
'The lion is eating John.'

(b) Eszi Jánost az irigység
eats J.-acc the envy-nom
'Envy is eating John.'

If Hungarian lacks the above subject/object asymmetry characterizing idiomaticity in configurational languages like English, then this provides support for a structural analysis wherein subject and object assume symmetric positions.

2.4. Movement of subjects

(iv) Subjects and objects in many constructions are extracted from their local clause with an equal ease in Hungarian. In English, the complementizer *that* blocks the extraction of the subject (a.k.a. the *that*-trace effect), whereas it has no effect on the extraction of the object (see (11)). Hungarian has no comparable *that*-trace effects (see (12)), hence subject-extraction behaves on a par with object-extraction in this regard. (Note that for many speakers the long-moved subject *wh*-expressions preferably appear in accusative case, licensed by the verb within the clause where they are moved to (see e.g., É. Kiss 1987b, (20)). No *that*-trace effect is attested in that variety either.)

(11) (a) Which candidate did you say (*that) became president?
(b) Which candidate did you say (that) the people elected?

(12) Melyik jelölt mondtad, hogy elnök lett?
which candidate-nom say-past-2sg that president become-past-3sg
'Which candidate did you say became president?'

Hungarian has no ban on *wh*-extraction across a local [Spec,CP] filled by another *wh*-phrase either. While in English the extraction of a *wh*-phrase across the local [Spec,CP] filled by a *wh*-expression leads to ungrammaticality with subjects, but only to a milder degradation with objects (cf. (13)), no such difference can be detected in Hungarian, cf. (14) (É. Kiss 1987b).

(13) (a) **Which boy do you wonder why wants to buy a new car?
(b) ?Which car do you wonder why John wants to buy?

(14) (a) (?)Melyik tanár nem tudod, hogy miért buktatott meg *pro*?
which teacher-nom not know-2sg that why failed-3sg pv (you.acc)
'**Which teacher do you wonder why flunked you?'

(b)[(?)]Melyik diákot nem tudod, hogy miért buktatták meg?
which student-acc not know-2sg that why failed-3pl pv
[?]'Which student do you wonder why they flunked?'

These two discrepancies between subject and object in English-type languages are normally accounted for in terms of the position of the subject. While the object is generated as sister to the verb (in a complement position), the subject is not: it originates (and also surfaces) higher. What É. Kiss concludes from the lack of these subject/object differences in Hungarian, is that not only the object, but also the subject is born as sister to the verb in a flat VP in this language.

2.5. Condition C

(v) Condition C (which requires referential expressions like names not to be c-commanded by a co-referential DP) rules out (15b), while it rules in (15a), because in English the subject c-commands the object, but the object does not c-command the subject.

(15) (a) Yesterday Peter$_i$'s mother phoned him$_i$
(b) *Yesterday he$_i$ phoned Peter$_i$'s mother

É. Kiss argues that in Hungarian Condition C effects obtain with R-expressions inside objects and subjects alike. (16) illustrates binding into the subject by the object.

(16)*Tegnap felhívta a fiúk$_i$ anyja$_{NOM}$ őket$_{ACC_i}$
yesterday up-called-3sg the boys-nom mother-poss.3sg-nom them
'Yesterday the boys'$_i$ mother called them$_i$ up.' (judgment from É. Kiss 2002)

This judgment once again follows from a non-configurational verb phrase structure, where the subject DP c-commands (into) the object DP, and vice versa.[6]

[6] Deletion rules also operate without a subject/object asymmetry, see (i). This has no bearing on the configurationality issue, however, since in these constructions the element escaping deletion (whether a subject or an object) is moved out of the ellipsis site prior to deletion (by focusing, topicalization or some other A-bar movement) (see É. Kiss 1994b; 2002).

2.6. Free postverbal consitutent order

(vi) The major descriptive appeal of the flat VP analysis is the ease with which it can treat the apparent radical freedom of postverbal word order. While we find a strict hierarchy to the left of the finite verb, in the postverbal area a radical freedom of constituent order is attested. This falls out in a flat VP analysis, on the assumption that the overt material to the right of the finite verb corresponds to what is dominated by the VP.

2.7. Anaphor and pronominal variable binding

(vii) S/O asymmetries are nevertheless manifested in anaphor and pronominal variable binding.

(17) (a) *Gyakran elemzi(k) önmaguk/egymás a pszichológusokat
often analyze-3sg/3pl themselves-nom/each other-nom the psychologists-acc
'*Themselves/each other often analyze psychologists.'

(b) Gyakran elemzik a pszichológusok önmagukat/egymást
often analyze-3pl the psychologists-nom themselves-acc/each other-acc
'Psychologists often analyze themselves/each other.'

These asymmetries are presumed not to be a reflection of a structural asymmetry, but instead, of an asymmetry in terms of prominence in a Thematic Hierarchy or in terms of linear precedence: É. Kiss (1991; 1994b) posits a Primacy Condition, which subsumes these two factors disjunctively in order to cover the complete set of anaphor and pronominal variable binding facts.[7]

(i) Marinak VIRÁGOT vett János, Zsuzsának pedig CSOKOLÁDÉT
M.-dat flower-acc bought J.-nom Zs.-dat as.for chocolate-acc
'As for Mary, John bought her FLOWERS, and as for Susan, he bought her CANDY.' (É. Kiss 1987b, (11))

(ii) Marinak JÁNOS vette a virágot, Zsuzsának pedig PÉTER
M.-dat J.-nom bought the flower-acc Zs.-dat as.for P.-nom
'As for Mary, JOHN bought her the flowers, and as for Susan, PETER bought her the flowers.'

[7] Notice that the flat VP structure causes Condition C to be violated in examples like (17b): the anaphor binds the referential expression within a flat VP. Condition B appears to be violated in a flat VP in examples like (i) below. In É. Kiss (2002), these unwelcome consequences are prevented by assuming the principle in (ii): since the subject DP can bind the thematically less prominent object DP,

All in all, the flat VP theory appears to be a descriptively successful and analytically simple account of the facts above taken together.

3. Reducing subject–object symmetries to scrambling

What I demonstrate next is that the arguments reviewed in the previous section are not compelling. The various forms of a lack of an S/O asymmetry **(i)**–**(vi)** enumerated in section **2** are inconclusive in supporting a non-configurational, flat VP approach. Arguments **(iii)**, **(iv)** and **(v)** are ill-founded, and **(i)**, **(ii)** and **(vi)** lose their force, given that a scrambling movement analysis based on a hierarchical *v*P can derive the observed patterns just as well.

3.1. Weak Crossover and Superiority

In particular, **(i)** and **(ii)** lose force because scrambling is known to obviate WCO violations. Scrambling languages typically lack WCO and Superiority violations in mono-clausal contexts (see (2) and (3), respectively) (e.g., Haider 1986; Saito 1992; Grohmann 1997; Wiltschko 1998; Fanselow 2001; compare Fanselow 2004). I illustrate this from German in (18) and (19), respectively.

---

the object cannot bind the subject, so Condition B and Condition C are in fact not disobeyed.

(i) Ő/*pro* megölte magát
he-nom/*pro*-nom pv-killed-3sg himself-acc
'He killed himself.'

(ii) **The asymmetry of binding**
If *a* can bind *b*, *b* cannot bind *a*.

Note that É. Kiss's Primacy Condition involves notions (relative thematic prominence, linear precedence) that are not directly available to build a grammatical analysis on within the current mainstream minimalist framework, where thematic roles are configurationally encoded in an articulated verb phrase structure (Hale–Keyser 1993b), and linear order is not encoded in syntactic structure (Kayne 1994). This Primacy Condition is actually akin to analogous principles of prominence utilized within the LFG framework in order to restrict binding relations. For instance, Bresnan's (1995; 1998) Prominence Principle involves a hierarchy of grammatical functions, precedence and thematic prominence; languages are then claimed to vary as to which of these constraints are active (which aspect of Bresnan's approach is, once again, not transposable to a minimalist model).

(18) Wen liebt seine Mutter nicht?
who-acc loves his mother-nom not
'Who is such that his own mother does not love him?'

(19) (a) Wen hat wer gesehen?
who-acc has who-nom seen
'Who saw whom?'

(b) Was hat wer gesehen?
what-acc has who-nom seen
'Who saw what?'

(c) Ich weiss wen wer liebt.
I know who-acc who-nom loves
'I know who loves whom.'

Scrambling languages are assumed to show no Superiority or WCO effects becasue scrambling itself obviates Superiority / WCO. (20) exemplifies WCO obviation in German, and the same is illustrated in (21) for Japanese:

(20) (a) *... weil seine Mutter jeden Studenten liebt
since his mother-nom every student-acc loves
'*... because his mother loves every student'

(b) ... weil [jeden Studenten]$_i$ seine Mutter t_i liebt
since every student-acc his mother-nom loves
(Grewendorf–Sabel 1999, 16)

(21) (a) ?*[[Soitui-no hahaoya]-ga [darei-o aisiteru]] no?
the-guy-gen mother-nom who-acc love Q

(b) ?Darei-o [[soitui-no hahaoya]-ga [t aisiteru]] no?
who-acc the-guy-gen mother-nom love Q
'?*Who does his mother love?' (Saito 1992, 73)

The obviation effect follows on the assumption that an object can undergo A-bar movement starting from a position **above** the subject, a position that is available to it precisely due to scrambling. As Fanselow (2001) points out, *was-für* split can strand the *für* + XP component of the complex *wh*-phrase in a scrambling position, providing evidence that scrambling can feed *wh*-movement in German, see (22a). This approach receives further confirmation from the fact that an *in situ* object *wh*-

phrase can overtly scramble above the subject *wh*-phrase, see (22b) (or over a non-specific indefinite subject, see Fanselow 2001, 414).[8]

(22) (a) Was hätte denn [*t* für Aufsätze] selbst Hubert nicht rezensieren wollen
what had prt [*t* for papers] even Hubert not review wanted
'What kind of paper would even Hubert not have wanted to review?'

(b) Wem hat was wer *t* gegeben?
who-dat has what-acc who-nom given
'Who gave what to whom?'

Although WCO S/O asymmetries are absent with short *wh*-movement and focusing, they obtain in some other cases (cf. Marácz 1989). I illustrate this in (23) with universal quantifiers. The contrast in (23) would be explained on a flat VP analysis by É. Kiss's (1991; 1994b) Primacy Condition on Binding involving thematic prominence and linear precedence disjunctively (see **(vii)** in section **2** above). But the contrast receives a straightforward explanation on a hierarchical *v*P account as well: covert Quantifier Raising (QR) of the universal QP (cf. Surányi 2002) produces a WCO configuration in (23a), but not in (23b).

[8] As far as Superiority violations are concerned, various other alternative analyses might in principle be applicable. For instance, Bošković contends in a series of papers (see Bošković 2002 and references therein) that if a functional head attracts (and enters an Agree relation with) multiple instances of the same feature, the attracted elements can move to the functional head in any order, given that the same total number of nodes will be crossed whatever the order of the movements. Based on work by Reinhart (1995; 1998) and Fox (1995; 1998; 2000) on what has come to be termed "interface economy" phenomena, another possible line is to argue that Superiority-violating multiple *wh*-fronting orders are licensed *qua* economy violations because they target an interpretation that cannot be achieved by the non-Superiority violating *wh*-order (an approach embraced in Fanselow 2004; see also Surányi 2002, ch. 6 for this point). Indeed the sorting keys (cf. Kuno 1982) in (6a) and (6b) above are different, and accordingly, appropriate answers differ too. (i) can answer (6a) but not (6b), and (ii) can answer (6b) but not (6a).

(i) [TOP János] [FOC tortát] csinált, [TOP Mari] [FOC jégkrémet], ...
J.-nom cake-acc made-3sg M.-nom ice-cream-acc,
'John made a cake, Mary made ice-cream, ...'

(ii) [TOP A tortát [FOC János] csinálta, [TOP a jégkrémet [FOC Mari], ...
the cake-acc J.-nom made-3sg the ice-cream-acc M.-nom
'The cake was made by John, the ice-cream by Mary, ...'

(23) Nem ismerte fel ...
not recognized-3sg pv

(a) *[az a férfi, aki bement pro_{3SGi} hozzá] [mindegyik lányt]
that the man-**nom** who in-went-3sg to.her$_i$ every girl-**acc**$_i$
'*The man who dropped by her$_i$ didn't recognize every girl$_i$.'

(b) [mindegyik lány]$_i$ [azt a férfit, aki bement pro_{3SGi} hozzá]
every girl-**nom**$_i$ that-acc the man-**acc** who in-went-3sg to.her$_i$
'Every girl$_i$ didn't recognize the man who dropped by her$_i$.'

Universal QPs (and other increasing distributive QPs) can be fronted in Hungarian to their preverbal scope-taking position overtly (traditionally identified as an adjunction site; Szabolcsi 1997 argues that it is DistP, but see Surányi 2002; 2004 for a defense of the traditional view). If we apply this overt form of QR (call it QP-fronting) in (23a), we get (24), and somewhat surprisingly, the degradation of (23a) almost completely disappears.

(24) (?)[Mindegyik lányt]$_i$ felismerte t_i' [az a férfi, aki bement
every girl-**acc**$_i$ recognized-3sg that the man-**nom** who in-went-3sg
pro_{3SGi} hozzá] t_i
to.her$_i$
'*?The man who dropped by her$_i$ recognized every girl$_i$.'

The Primacy Condition, which disjunctively involves precedence and thematic prominence as a condition on binding, is able to cover this fact: the quantifier precedes the bound pronominal in (24).[9] However, the same fact receives an explanation on the scrambling account too, and does so in the same way as in the case of (3a) above: in the derivation of (24), the object is first scrambled to a position above the subject ($= t_i'$), and is A-bar-moved to its preverbal position in a second step. What explains that

[9] However, the disjunctive definition runs into a problem with simple cases like (i).
(i) *Szereti Jánost önmaga
love-3sg J-acc himself-nom
If binding of B by A is licensed either if A thematically more prominent than B, or if A precedes B, then (i) is predicted to be grammatical, contrary to fact. If, however, only thematic prominence matters, but linear precedence does not (cf. É. Kiss 2002), then the apparent feeding effect of placing the object to the left of the subject on A-binding of anaphors and pronominal variables, which I analyze here as an effect of scrambling, and which is discussed extensively in É. Kiss (1991; 1994b) (see also (7), as well as section 6 below for examples of this), is left without any account.

this derivation is not available in (23a), is that scrambling is generally restricted to overt syntax.[10]

In short, on a scrambling account, thematic prominence can be replaced with c-command inside the *v*P, and instead of precedence, scrambling takes care of the availability of A-binding by the object into the subject precisely when the object comes to precede the subject. On this approach, the licensing condition of binding can simply be based on c-command, instead of the theoretically less desirable disjunctive principle of the Primacy Condition.

3.2. Idioms and compositional theta-role assignment

The appealing argument from idioms and compositional theta-role assignment **(iii)** is inconclusive for two reasons (for the sake of brevity, I

[10] É. Kiss (1994a) notes that Brody (1990) discusses the example below, which he marks as *?. É. Kiss claims that context can improve it into a grammatical, though still degraded, sentence (namely, if the universal QP quantifies over a familiar and salient set). These judgments fall into place under the present view. If the set quantified over by the universal quantifier is familiar/salient, then it can (somewhat marginally) function as an information structural topic. In this case, on its way to the left peripheral landing site, the QP can touch down in a scrambled position, which explains (23b). If, however, these conditions are not met, then a universal QP like *mindenkit* 'everyone' is difficult to be construed as a topic, it will normally function instead as the information focus of the sentence (in this case the immediately following complex verb can undergo stress eradication). This discourse structural status does not allow the QP to undergo scrambling prior to QP-fronting, whence it is fronted to the left periphery in one step, giving rise to a WCO effect.

(i) *?Mindenkit felismert a férfi, aki belépett a szobájába
everyone-acc pv-recognized the man-nom who in-stepped the (her) room-in
'*The man who entered her room recognized every girl.'
(judgment Brody's)

Dobrovie-Sorin (1990) points out an analogous contrast in Romanian in the domain of *wh*-elements: whereas D-linked object *wh*-phrases can escape inducing a WCO violation, non-D-linked ones cannot. It is interesting in the present context to note a recently discovered parallel in English. Ishii (2006) shows that object *which*-phrases (i.e., overtly D-linked *wh*-expressions) in English fail to show a WCO effect in their local clause, which he ascribes to the movement step that targets the *v*P-edge (performed by *wh*-objects in the course of their successive cyclic movement to CP). Given that this intermediate step (which is likened by Ishii to scrambling) may count as an A-movement operation only in the case of D-linked *wh*-phrases, the obviation effect of this movement operation is limited to *which*-phrases.

concentrate here on idioms, but the arguments extend also to compositional theta-role assignment). First, the logic of the argument is flawed: on a flat VP analysis, which the evidence is supposed to support, [V + S] does not form a base structure constituent, and nor does [V + O]. This apparently flies in the face of the notion (going back to Marantz 1984) that idioms are (roughly) base structure constituents. Second, a number of idioms cited to instantiate the [V + S] idiom type are not in fact disallowed in a hierarchical VP structure on Marantz's (1984) assumptions either. For instance, [V + S] idioms involving a subject that is arguably an underlying internal argument of the verb, such as unaccusatives, are predicted to be allowed. *Piroskának leesett az álla* lit. 'Piroska-dat fell the jaw' and *Piroskának kinyílt a szeme* lit. 'Piroska-dat opened the eye' (cited in É. Kiss 2002) exemplify this type of idiom.[11] Psych verb constructions are another case in point. Chtareva (2005) argues that a group of [V + S] idioms in Russian that are apparently problematic for Marantz's (1984) hypothesis in reality fully conform to it, insofar as they represent idioms involving psychological causative predicates whose surface subjects are themes, and whose surface objects are experiencers (like one of the two uses of the English verb *frighten*); see (8a). On a fairly standard approach (see Belletti–Rizzi 1988), in these constructions the experiencer is generated above the theme, hence the verb and the surface subject form a base structure constituent.[12] This type of example has often been used, albeit wrongly, to back up the flat VP analysis (see e.g., É. Kiss 1987b, 22–23); see (25b), as well as (8) above.

(25) (a) Ivan-a zajela sovest'
Ivan-acc up.ate conscience-nom
'Ivan's conscience was troubling him.'

(b) Jánost elkapta a gépszíj
J-acc pv-caught the driving.belt-nom
'John is intensively involved/caught up in something.' (É. Kiss 2003, 26)

[11] These two examples involve a possessor that has been extracted out of the underlying complement DP (cf. Szabolcsi 1983). Idioms with an open possessor slot are possible, independently (e.g., *cat got* x*'s tongue*), as the possessor is not an argument of the verb.

[12] Nunberg et al. (1994) argue that many idioms are in fact compositional: the parts of these idioms have contextually restricted metaphorical interpretations, which combine transparently (see also Marantz 1997). This is what happens in examples like (25a,b), too.

Third, even English has idioms involving S and V, but not the complement of V: for instance, *God bless him, Fortune smiled on Gwendolyn* or *The devil alone knows X* (see Postal 2002 for more examples, and compare also Everaert 1993; the same goes for the influence on theta role assignment, e.g., *Somebody is eating popcorn* vs. *Something is eating him*). According to Everaert (*ibid.*), subject idioms are much more frequent in clearly configurational languages than often suggested in the literature, although object idioms are clearly the less marked case.[13] Of course, the same is true for Hungarian: [V + O] idioms are more abundant than [V + S] idioms. All in all, no firm conclusions can be drawn from the domain of idioms to back up a flat VP analysis.

3.3. Movement of subjects

Moving on to the observations in **(iv)**, it is easy to see that, even though they involve a difference between subject and object, they are quite independent of the issue of (non)configurationality. As far as the lack of *that*-trace effects is concerned, this has been correlated with the property of *pro*-drop (Perlmutter 1971), and with the availability of *v*P-internal subjects (Bennis 1986; Szczegielniak 1999), properties that are applicable to Hungarian and that can be found in configurational languages as well.

As for the general availability of *wh*-extraction across a local filled [Spec,CP], this is a feature that can be put down to the left-peripheral multiple specifier configuration underlying Bulgarian-type multiple *wh*-fronting (Rudin 1988), which is also found in Hungarian (see Surányi 2006). It has also been suggested that this behavior is a feature of languages where a *v*P-internal surface position is available to subjects, e.g., Italian, Spanish (see Sabel 2002 and references therein), which is once again a property that apparently holds of Hungarian. The availability of a *v*P-internal position for the subject once again does not directly concern the hierarchical asymmetry between the position of the subject and that of the object.

3.4. Condition C

Let us now come to the alleged S/O symmetry with respect to Condition C violations, i.e., **(v)**. The first point I would like to make concerns the

[13] See Martin Everaert's clarificatory note on the *Linguist List*, Vol-4-122.

status of examples like (16). 10 out of the 25 informants whose judgments I have had access to found examples analogous to (16) degraded, but not unacceptable (? or ??), and 7 speakers judged them to be OK, and only 8 informants rejected them as ?* or *. Second, the degradation found in (16) can partly be put down to the placement of the pronoun, which is in a final position, separated from the verb by the subject phrase. In Hungarian such a surface position is known to be generally disfavored by personal pronouns, which, if postverbal, prefer to be close to the verb (Varga 1981), not separated from it by a stress-bearing element. Indeed when the subject expression is fronted to a topic position and hence the accusative pronoun follows the verb immediately, the judgment profile improves significantly: OK = 10, ? = 9, ?? = 3, ?* = 2, * = 1. A more radical improvement is attested when the antecedent of the pronoun is made salient by the context, and the (3sg) overt pronoun in examples analogous to (16) is replaced by a (3sg) object *pro*. In contrast, when the silent pronoun is a subject bound by the possessor in the object, the sentence is severely degraded.[14]

[14] A similarly strong contrast is found with overt oblique case-marked internal argument pronouns, which lack a covert counterpart, see (i–ii). (More precisely, oblique pronominal expressions are realized as an element corresponding to the oblique case marker, whose morphosyntactic form is that of a possessed noun head, and whose possessor is the personal pronoun itself, typically a *pro*).

(i) ?A legjobb barátom$_i$ anyja gyerekként egyáltalán nem foglalkozott vele$_i$
the best friend-poss.1sg mother-nom child-as at.all not took.care with.him
'My best friend's$_i$ mother didn't take care of him$_i$ as a child at all.'

(ii) *A legjobb barátom$_i$ anyjával idős korában nem foglalkozott ő$_i$ egyáltalán
the best friend-poss.1sg mother-with old age-poss.3sg-in not took.care he at.all
'*He$_i$ did not take care of my best friend's$_i$ mother in her old age.'

It must be noted that prosodic context seems to affect the acceptability level of sentences like (16): when followed by a stress-bearing element (as in (iii)), the acceptability profile of sentences like (16) involving an overt object pronoun improves noticeably. As has been noted in the main text, it also enhances acceptability if the object pronoun is not separated from the verb by a stress-bearing element. Using a dative possessor instead of the nominative form is another factor that increases acceptability for some speakers. Another improving factor according to the judgment of some informants is the topicalization of the subject DP containing the antecedent possessor. By contrast, none of these factors alter the judgment of overt subject pronouns.

(iii) ?(?)Hét közben már nem is hívja fel a fiúk anyja
week during anymore not also calls up the boys-nom mother-poss.3sg-nom
őket egyáltalán
them at.all
'The boys' mother does not call them anymore at all during the week.'

(26) (a) ?Péter$_i$ főnöke hívta fel *pro*$_i$ reply to: Who called up Peter?
Peter's$_i$ boss-nom called up him$_i$

(b) *Péter$_i$ főnökét hívta fel *pro*$_i$ reply to: Who did Peter call up?
Peter's$_i$ boss-acc called up he$_i$

It is important to note that although examples with an object pronoun co-referring with a lexical possessor inside the subject are of varied acceptability across speakers, speakers tend to find sentences with a subject pronoun co-referring with a lexical possessor inside the object much worse. Although judgments of co-reference (involving nominative and accusative pronouns—but see note 14) are not so sharp as to serve as the basis of a strong argument either pro or con, they lean in the expected direction only if the subject is indeed generated above the object.[15]

[15] As for the reason for the preference of personal pronouns to surface immediately after the verb (or verb plus particle), it can be speculated that this is due to their prosodic properties and/or the familiarity of their referents. On either account, it may well be that they preferably undergo scrambling (to the right of the verb).

If this is correct, then it makes available two potential ways to capture why sentences of type (16) are degraded. One possibility is to construe (16) as involving the scrambling of both S and O (call this 'dual scrambling'): in this case the scrambled object pronoun will bind the base copy of the R-expression inside the subject. The fact that a scrambled object anaphor does not give rise to ungrammaticality even though it c-commands the base copy of its antecedent R-expression in seeming violation of Condition C does not in intself speak against such a "dual scrambling" analysis of the degradation of (16). This is because this particular behavior of object anaphors in relation to the subject R-expression is well-known to be an independent property in scrambling languages as different as Japanese, Hindi, Korean, German and Georgian (a property put down to "lethal ambiguity" by McGinnis 2004); see (40) for the Hungarian example. What argues against this account, however, is that object personal pronouns are known not to be exempt, in the manner object anaphors are, from inducing a Condition C violation in the very same language type, as they are not in Hungarian either (see (i) below).

(i) *Tegnap felhívta [őket]$_i$ [a fiúk$_i$ anyja] *t*
yesterday up-called-3sg them the boys-nom mother-nom
'Yesterday the boys' mother called them up.'

According to the second possibility, (16) is construed as not involving scrambling at all, and its degradation is therefore due to the above-mentioned preference of personal pronouns to undergo scrambling, which the object pronoun in (16) fails to satisfy. Scrambling the object pronoun to the left of the subject in (16) cannot help either: in such a scenario the object personal pronoun binds the R-expression within the subject from its scrambled position, inducing a Condition C violation (see (i)).

Interestingly, É. Kiss has suggested that with *wh*-possessors (instead of lexical possessors) we get no S/O asymmetry, unlike in English (judgments from É. Kiss 1987a), compare (27) and (28). On the flat VP approach, (27) can be seen as involving Strong Crossover, i.e., a Principle C violation.

(27) (a) *Kinek$_i$ az anyja hívta fel őt$_i$? (cf. (28a))
whose$_i$ the mother-poss.3sg-nom called up him$_i$

(b) *Kinek$_i$ az anyját hívta fel ő$_i$? (cf. (28b))
whose$_i$ the mother-poss.3sg-acc called up he$_i$

(28) (a) Whose mother called him up?
(b) *Who did his mother call up?

Let us accept É. Kiss's judgments in (27) at face value (though, see note 16 for qualification). What I would like to argue is that even given these judgments, such a subject/object symmetry does not necessarily provide evidence for a flat VP analysis. The ungrammaticality of (27b) follows if S c-commands O inside the *v*P. As for (27a), I propose that it is ruled out because it is blocked by (29).

(29) Kit$_i$ hívott fel t_i' az (ő$_i$) anyja t_i? (=(3a))
who-acc called up the his mother-poss.3sg-nom
'*?Who$_i$ did his$_i$ mother call up?'

Ruys (1994) argues that, given an interface economy approach (see Reinhart 2006 and references therein), (28b) is blocked in English by (28a), because (28a) is derivationally more economical than (28b), as it involves a shorter *wh*-movement (cf. also Spector 2004). On account of its optionality, scrambling is often taken to incur no derivational cost (see e.g., Fukui 1993; Saito–Fukui 1998; Boškovič–Takahashi 1998; note that this follows directly on a base-generation account of scrambling). Recall that I assumed in section **3.1** above that (29) (=(3a)) is well-formed in the first place because of the availability of a derivation involving scrambling of O *above* S prior to *wh*-movement, i.e., to a position that is closer to the left peripheral CP/FocP position than the base position of S. This means that the derivation of (29) involves a shorter *wh*-movement to CP/FocP than (27a), and I propose that this is why (27a) is blocked.

An analogous paradigm is found with universal quantifiers in the place of *wh*-phrases, and the same blocking effect will be triggered. I

omit the examples here in the interest of saving space.[16] Note finally that (27a) is also out in German. German is configurational and has scrambling, hence the same logic of blocking applies there as well.

3.5. Free postverbal constituent order and verb raising

The freedom of postverbal constituent order, i.e., **(vi)**, is clearly not compelling evidence in favor of a flat VP insofar as scrambling can derive the freedom in word order just as well. Scrambling is predicted to be restricted to the postverbal field, once it is assumed that the verb is moved to the head of a functional projection above the *v*P. That the verb is raised into the IP domain (in neutral sentences) is a view shared among

[16] In a late lexical insertion model like Distributed Morphology, one can have (27a) and (29) stem from the same Numeration, if one makes the assumption that *wh*-pronouns in Hungarian are nothing else but (spellout forms corresponding to) pronouns in the local context of a *wh*-feature (either on D or on Foc); this matter will not be pursued here. In fact, Ruys's (1994) conception of blocking in (28) derives from the view that the competing (reference) set of derivations is determined by interpretive equivalence, rather than on a Numeration of lexical items (see also Fox 2000 and Reinhart 2006 and references therein; the special relevance of Fox's implementation of this view is that his account is formulated in terms of the relative length of movement paths).

The interface economy approach is supported by the fact that if the *wh*-element *ki(nek)* 'who(se)' in (27a) is replaced by the a D-linked *wh*-phrase like 'which boy,' then (27a) improves significantly, see (i). (In comparison, performing the same replacement in (27b) does not result in any improvement.) By the same logic of blocking as applied above, (i) should be degraded just as much as (27a) is, given that there exists a more economical derivation targeting the same interpretation (involving object scrambling prior to *wh*-movement), see (ii). The reason why the same logic is inapplicable to (i–ii), I believe, lies in the fact that (i) and (ii) are not entirely synonymous: informally, while (i) is a question about a set of mothers (as a function of a set of boys), the question in (ii) quantifies directly over a set of boys. The non-identity of the LF representations of (i) and (ii) actually follows on Rizzi's (2001) theory of A-bar reconstruction, whereby only non-D-linked *wh*-expressions have their descriptive restriction obligatorily reconstructed, whereas the same is not enforced in the case of D-linked *wh*-phrases, whose descriptive restriction is contextually given, topic-like, and as such they can remain in the left periphery, licensed there as topics generally are (cf. also Heycock 1995). Thus, the lexical restriction undergoes reconstruction in (27a), yielding the same LF representation as (29), which can be informally given as ?x. x*'s mother called* x. In contrast, the lexical restriction does not necessarily reconstruct in (i) (see Shavrit–Guerzoni 2003 for an argument for the stronger view that it cannot), therefore it can (or, following Shavrit – Guerzoni *ibid.*, it must) produce an LF representation different from that of (ii).

others by Szabolcsi (1997), Puskás (2000), and Brody–Szabolcsi (2003). The exact identity of the projection hosting the verb is immaterial for the present purposes. Determining the exact landing site (and potentially, also a trigger) of the Hungarian scrambling movement operation is tangential to the main point of the present paper, and indeed the choice is underdetermined by the data discussed in these pages (e.g., scrambling targeting the *v*P-edge, or the TP-edge are equally conceivable, depending, of course, on the choice of specific theoretical assumptions;[17] for recent alternatives, see e.g., Boškovič–Takahashi 1998; Grewendorf–Sabel 1999;

(i) ?(A három közül) melyik fiúnak$_i$ az anyja hívta fel
the three out.of which boy-dat the mother-poss.3sg-nom called up
őt$_i$ idejében?
him time.in
'Out of the three boys, which boy's mother called him in time?'

(ii) (A három közül) melyik fiút$_i$ hívta fel az (ő)$_i$
the three out.of which boy-acc called up the he
anyja idejében?
mother-poss.3sg-nom time.in
'Out of the three boys, which boy did his mother call in time?'

To the extent that one can interpret *kinek az anyja* 'whose mother' in (27a) as D-linked in a given context, the same processes that I have argued to apply in (i) can—to some degree—mitigate the unacceptibility of (27a) (here the descriptive restriction, besides *person*, is derived from the discourse context). Indeed, a number of speakers that I have consulted find (27a) marginally acceptable (once again, the factors discussed in section **3.4** in relation to (16) apply to (27a) as well, to the relative improvement of its acceptability).

Note that the present account of (27a) presupposes that the subject cannot raise to a srambled position: otherwise the *wh*-movement of the subject in (27a) and the *wh*-movement of the scrambled *wh*-object in (29) could be equally short, in which case (29) could not block (27a). That local subject scrambling is unavailable is argued (for Japanese) by Saito (1985), and is a reasonably well-established generalization in the literature on Japanese-type scrambling (see Ko 2005 for an argument for an opposing view).

[17] Note that if scrambling targets the *v*P-edge, say, by adjunction to *v*P, then the blocking analysis of (27a) and (29) is compatible only with such a metric of the length of movement paths that is sensitive only to categories that *properly contain* the moved element at its pre-movement position. Such a metric determines the movement of a scrambled object (as in (29)) to be shorter than the movement of a subject out of a *v*P where no object scrambling has taken place (as in (27a)): the *v*P category is contained in the movement path only in the latter case. If the target of scrambling is (exclusively) the *vP*-edge, then it must be

Karimi 2003; Kitahara 2002; Miyagawa 1997; 2001; 2003; Saito 2003). Therefore, the issue is not discussed here in any detail.[18]

3.6. A-binding

Finally, the A-binding S/O asymmetries (= **(vii)**) can be captured in a hierarchical *v*P without directly relying on thematic prominence or linear precedence, or indeed a disjunctive definition incorporating both: A-binding facts can be deduced from structural asymmetries in the hierarchical structure in terms of c-command. The issue of A-binding will be taken up and will be dealt with in more detail in section **5**.[19]

Having shown that some of the arguments for a flat VP are ill-founded, and others are forceless once a scrambling account is considered as an alternative, in the next section I go on to present phenomena of S/O asymmetries that seriously challenge the non-configurational VP analysis, and directly bolster a scrambling approach (*modulo* a hierarchical *v*P).

---

ensured that adjuncts can intervene between a scrambled phrase and the subject in [Spec,*v*P] (because an *Object* > *Adjunct* > *Subject* postverbal order is well-formed in Hungarian). Neither of these two conditions applies to an approach that takes scrambling to target the TP-edge (or allows scrambling to target either the *v*P-edge or the TP-edge). This latter account presupposes that the verb in a neutral clause sits in a functional projection even higher than the TP (say, in the head of the projection whose specifier is occupied by the verbal particle in a neutral sentence). An advantage of the former view, however, is that it can straightforwardly account for the unavailability of subject scrambling (see the previous note), which would be ruled out *qua* vacuous movement taking place **within** the edge of a projection (viz. *v*P).

[18] Although it apparently provides a simple account of postverbal free word order, adopting a flat VP implies giving up the binarity of Merge, and it is also diametrically opposed to what Kayne's Linear Correspondence Axiom (LCA) permits (one of the consequences of the LCA is that "if two phrases differ in linear order, they must also differ in hierarchical structure", Kayne 1994, 3). É. Kiss (2002) speculates that "a relaxed version" of Kayne's (1994) LCA is not necessarily at odds with a flat VP: elements under VP are unordered precisely because they do not asymmetrically c-command each other. This "relaxed version" is not provided, however. Should the LCA be relaxed in such a way as to allow structures involving symmetric c-command, as in the case of a flat VP, a whole range of welcome results of LCA-based syntax would be effectively lost (among others, the way aspects of X-bar theory are derived by the LCA).

[19] As Surányi (in press) shows in some detail, Hungarian does not share the properties of either one of the two well-studied classes of non-configurational languages. This further weakens the position of the flat VP approach.

4. Arguments in favor of the hierarchical *v*P + scrambling account

In addition to the S/O asymmetry exhibited by universal QPs for WCO, which was discussed in section **3**, in this section I point out several other S/O asymmetries. These asymmetries are all problematic for a non-configurational VP approach, but are expected if the Hungarian *v*P is hierarchical.[20]

4.1. Superiority

The first asymmetry to be noted here concerns effects of Superiority, which do obtain in various constructions. The illustrative example in (30) involves *n*-word fronting, where obviation by scrambling (cf. section **3**) is dispreferred. Scrambling is disfavored (as an intermediate movement step) in the derivation of (30) due to the fact that the discourse effect that scrambling results in, i.e., familiarity, is incompatible with the non-specific (non-familiar) interpretation of the fronted object *n*-word in (30). In a context, however, where the object *n*-word can be interpreted as specific (quantifying over a familiar set), the pattern in (30) becomes acceptable. As expected under a configurational analysis of *v*P, if the subject *n*-word is fronted instead of the object *n*-word in (30), the sentence is acceptable once again.

(30) #Nyilvános helyen szerintem senkit se csókoljon meg senki
public place-on in.my.view noone-**acc** neg kiss-imp-3sg pv noone-**nom**
intended: 'I think nobody should kiss anybody in a public place.'

The same holds true of multiple *wh*-questions in which one *wh*-element is fronted only, asking for a single pair of individuals, in particular, of the type that involves two non-D-linked *wh*-expressions. In the dialogue below, the inspector (I) can ask the witness (W) the question in (a), while question (b) is infelicitous, given that neither of the two *wh*-pronouns is D-linked.

(31) W: I heard the noise of someone slapping someone else in the face behind my back. I turned around at once.
I: And what did you see?

[20] These asymmetries are not covered by the Primacy Condition on binding (cf. **(vii)** in section **2**).

(a) Ki vágott pofon kit?
who-nom hit-past-3sg face.on who-acc
'Who slapped whom in the face?'

(b) #Kit vágott pofon ki?

Scrambling is unavailable to the non-D-linked *wh*-object, whence it can only move to the left peripheral CP/FocP from its VP-internal position. This, however, results in a Superiority effect.

4.2. Movement out of subjects

A second difference between S and O, one that is expected on a configurational analysis of *v*P, is that subjects, but not objects (and other complements) are CED islands, similarly to what we find in English. If not only objects, but subjects are also complements of the verb, as the flat VP account presumes, then such asymmetries are unexpected.

(32) (a) ?Melyik tisztviselővel$_i$ olvastál [egy interjút t_i]?
which official-with read-past-2sg an interview-acc
'Which official did you read an interview with?'

(b) *Melyik tisztviselővel$_i$ állította [egy interjú t_i], hogy nő a GDP?
which official-with claimed an interview that grows the GDP
'With which official did [an interview *t*] claim that the GDP is growing?'

4.3. Condition C

Although judgments go in the direction expected on a configurational *v*P account, Condition C effects involving overt nominative and accusative pronouns do not result in a very sharp contrast between S and O, as discussed in section **3.4** (perhaps due to factors discussed there, see especially note 14). However, Condition C effects do produce a strong S/O asymmetry in the domain of epithets, i.e., definite NPs which are coreferential with, though different in descriptive content from, their antecedent; see (33a–b). These function like pronouns, but can be used for testing purposes here free of the complications associated with pronouns (again, cf. section **3.4**). Similarly, we find a marked S/O asymmetry for Condition C in A-bar reconstruction (33c–d), and with lexical DPs (33e–f) (the latter is noted by Marácz 1989, and by Choe 1989). In (33c–d) the object and the subject, respectively, are fronted to the topic position. This

A-bar movement is reconstructed to the position marked by the trace (Chomsky 1993; 1995) i.e., to a *v*P-internal position.

(33) (a) János anyja$_k$ nem is látogatja t_k azt a szerencsétlen gyereket
John's$_i$ mother-**nom** not even visit-3sg that-acc the poor child-acc$_i$
'John's$_i$ mother does not even visit that poor child$_i$.'

(b) *Az a szerencsétlen gyerek$_k$ nem is látogatja t_k János anyját
that the poor child-**nom**$_i$ not even visit-3sg John's$_i$ mother-acc
'*That poor child$_i$ does not even visit John's$_i$ mother.'

(c) *[A Jánossal$_i$ való beszélgetésünket]$_k$ később letagadta (ő$_i$) t_k
the J$_i$-with expl discussion-poss.1pl-acc later pv-denied-3sg he$_i$
'*He$_i$ later denied our discussion with John$_i$.'

(d) [A Jánossal$_i$ való beszélgetésünk]$_k$ rossz színben t_k tüntette fel őt$_i$
the J$_i$-with expl discussion-poss.1pl-nom bad color-in showed pv him$_i$
'Our discussion with John$_i$ gave him$_i$ a bad reputation.'

(e) Felhívta János$_i$ anyósa Jánost$_i$
pv-called-3sg J's$_i$ mother.in.law-**nom** J-acc$_i$
'John's$_i$ mother-in-law called John$_i$.' (adapted from Marácz 1989)

(f) *Felhívta János$_i$ János$_i$ anyósát.
pv-called-3sg J-**nom**$_i$ John's$_i$ mother.in.law-**acc**
'*John$_i$ called John's$_i$ mother-in-law.'

These data demonstrate that Condition C does in fact tease apart subject from object, as far as their base positions are concerned: the subject in Hungarian too originates higher than the object. The same conclusion is suggested by the observation (illustrated in note 14 above) that while various factors (namely, prosody, topicalization of the DP containing the antecedent possessor, and the case-form of the possessor) influence the acceptability of an object pronoun coreferring with the possessor inside the subject, the same factors do not affect the (non-)acceptability of a subject pronoun coreferring with the possessor contained in the object. This latter fact is predicted on the scrambling account, as only the latter scenario involves a Condition C violation, given a hierarchical *v*P.[21]

[21] Returning to the examples with covert (*pro*) pronouns in (26), it is conceivable that they do not involve binding per se, in which case (26) is analogous to (33e–f) above (cf. Reinhart 1983). Without a context, (26a) is strongly ungrammatical. The possessor in (26) is apparently not salient enough in itself, i.e., without a context, to license a *pro*, which is known to require a highly salient antecedent. This is confirmed by the examples below, where the *pro* element is one clause down from the possessor, whence Condition C cannot explain why these sentences are out.

4.4. Scope-taking of non-increasing QPs

Another domain where an S/O asymmetry is detected is scope-taking by postverbal non-increasing QPs (increasing QPs take scope via a mechanism distinct from that involved in scope-taking by non-increasing QPs, see Szabolcsi 1997 and Surányi 2004 for diverging views). A *few*-QP_{OBJ} cannot scope over a uQP_{SUBJ} (34a), while a *few*-QP_{SUBJ} can scope over the uQP_{OBJ} (34b):

(i) *Mari$_i$ anyjának elmondtam, hogy nem kedvel *pro*$_i$ már engem
M.-nom mother-dat pv-tell-past-1sg that not like-3sg (she) anymore me
'I told Mary's mother that she doesn't like me anymore.'

(ii) *Mari$_i$ anyjának elmondtam, hogy nem kedvelem *pro*$_i$ már
M.-nom mother-dat pv-tell-past-1sg that not like-1sg (her) anymore
'I told Mary's mother that I don't like her anymore.'

Oblique pronominals seem to tolerate an antecedent of the degree of salience associated with a possessor, as witnessed by (i) of Note 14 above. Following this line of thought, a potential explanation for the degradedness of É. Kiss's example (16) as well as that of (27a) could be based on the requirement of the degree of salience (accessibility) imposed by the overt third person pronoun on its antecedent. The degradedness of (26) may then follow, insofar as a possessor in the subject is simply not salient enough to serve as an antecedent of an overt third person object pronoun either. Indeed, if *pro* in (i) is replaced with an overt pronoun, the acceptability of the sentence does not significantly improve. The accessibility requirements of anaphoric forms are known to vary (see e.g., Ariel 1994); this might be the reason underlying the fact that the Hungarian data involving overt personal pronouns (in object position) differ from their English counterparts. That it is not c-command, but salience that is at issue in (16) is also compatible with the observations based on examples like (33a–b): it is well-established that epithets impose a different requirement of salience than third person pronouns. Another factor that matters for salience is the level of embedding. The more deeply the antecedent is embedded, the less salient it is. Thus it is expected that the overt third person object pronoun will be able to take as its antecedent a non-possessor nominal inside the subject, as in (33d). The fact that for some speakers dative possessors inside the subject are better licensors of coreferential object pronouns than nominative ones can also be made sense of in the very same terms, given that the dative possessor is known to occupy a higher (in fact, a left-peripheral) position within the DP than their nominative counterpart. Topicalization of the subject DP, which is another improving factor in the licensing of the object pronoun (see Note 14) can be also explained in terms of accessibility: topicalization enhances the salience of the antecedent possessor. I have not been able to study the salience requirements of the various pronominal forms in sufficient detail, therefore these considerations remain tentative, and will not be pursued here any further.

(34) TAVALY végzett el ...
last.year did-3sg pv ...

(a) minden diák kevés kurzust (S > O, *O > S)
every student-**nom** few course-**acc**
'It was last year that every student did few courses.'

(b) kevés diák minden kurzust (S > O, O > S)
few student-**nom** every course-**acc**
'It was last year that fewer than 100 students did every course.'

This is because decreasing QPs do not take inverse scope higher than their A-position (see Szabolcsi 1997 and Surányi 2004 for detailed discussion and references). The contrast in (34) is explained only of the A-position of the subject is higher than the A-position of the object.[22]

4.5. Incorporation

As Marácz (1989) points out, incorporation of a bare nominal is possible when the nominal is an object, but impossible when it is a subject. This is exactly what is predicted in Baker's (1988) model of incorporation as involving syntactic (upward) head-movement, provided, of course, that the subject is generated higher than the object.

(35) (a) János könyvet olvas
J.-nom book-acc read-3sg
'John is reading a book.'

[22] As for increasing quantifiers like universal QPs, their relative scope in the postverbal domain is known to be free with respect to each other (e.g., É. Kiss 2002) (even though stress seems to influence relative scope for many speakers). This situation is not different from that of languages like English, where verb phrase internal increasing quantifiers can also take both wide and narrow scope with respect to each other (with some exceptions, like the double object construction, which lacks a direct counterpart in Hungarian). This basic fact of English is conveniently captured in a standard Quantifier Raising based approach, and I proposed in Surányi (2002; 2004) to apply a QR-based analysis to Hungarian as well (*contra* Szabolcsi 1997). As for focused elements, they also exhibit freedom of relative scope within the postverbal domain, a generalization that I argued to capture in terms of covert focus movement in Surányi (2002; 2004). In short, the apparent lack of syntactic restrictions of postverbal relative scope does not bear on the issue of the configurationality of the verb phrase.

(b) *Tanár olvas egy jó könyvet
teacher-nom read a good book-acc
'Teacher(s) read(s) a good book.'

Thus far I have presented arguments in favor of the approach that incorporates a hierarchical VP (i.e., *v*P) structure and postverbal scrambling (N.B. the verb overtly raises out of the *v*P). In the remainder of the paper I demonstrate that the reordering of the object to the left of the subject in the postverbal field indeed has the properties of (a certain type of) scrambling movement.[23]

5. Probing the properties of Hungarian scrambling

If Hungarian indeed has a configurational *v*P, with the subject generated higher than the object, and postverbal object–subject order is indeed the result of scrambling movement, we expect sentences with this order to exhibit properties normally displayed by scrambling orders in well-known scrambling languages. Given that several distinct types of scrambling languages and scrambling operation types have been described in the literature (cf. e.g., the German-type vs. Slavic-type vs. Japanese type oppositions), probing the properties of what I have assumed to be a scrambling movement will also involve situating Hungarian scrambling (descriptively) within the scrambling typology.

5.1. Scrambling and anaphor binding

Scrambling of the object above the subject feeds the binding of anaphors in the possessor position of the subject in Hungarian (see 36a–b).

(36) (a) *?Sokat kritizálják egymás szülei Jánost és Pétert
lot-acc criticize-3pl each other's parents-nom J.-acc and P.-acc

[23] It has also been pointed out for Hungarian (e.g., Speas 1990) that PRO in non-finite clauses can only function as a subject, but not as an object. This follows on theories of control where the syntactic position of the subject is different from that of the object. Note, however, that this does not necessarily turn into an argument in favor of a hierarchical verb phrase, as the position at issue is that of the verb phrase external, canonical subject position. Nevertheless, it still needs to be ensured on a flat VP approach that only a verb phrase internal subject, but not a verb phrase internal object can move here.

(b) ?Sokat kritizálják [Jánost és Pétert]$_i$ egymás szülei t_i
lot-acc criticize-3pl J.-acc and P.-acc each other's parents-nom
'John and Peter are criticized a lot by each other's parents.'

This property is characteristic of Japanese local scrambling (cf. (37) below, see Saito 1992, 74f); whereas it is not shared by German, Slavic or Albanian scrambling (see, e.g., Grewendorf–Sabel 1999; Kitahara 2002; Saito 2003; Karimi 2003, and references therein). (38) exemplifies the case of German.

(37) (a) ?*[[Otagai-no sensei]-ga karera-o hihansita] (koto)
each other-gen teacher-nom they-acc criticized (fact)

(b) ?[Karera-o$_i$ [[otagai-no sensei]-ga t_i hihansita]] (koto)
they-acc each other-gen teacher-nom criticized (fact)
'*Each other's teachers criticized them.'

(38) (a) *... weil [die Lehrer von sich$_i$] zweifellos den Studenten$_i$
since the teachers-nom of *sich*] undoubtedly the student-acc
in guter Erinnerung behalten haben
in good memory kept have
'The teachers of himself have undoubtedly kept the student in good memory.'

(b) *... weil [den Studenten]$_i$ [die Lehrer von sich$_i$] zweifellos t_i
since the student-acc the teachers-nom of *sich* undoubtedly
in guter Erinnerung behalten haben
in good memory kept have (Grewendorf–Sabel 1999)

This follows if Hungarian scrambling is or can be A-movement and Condition A is an "anywhere condition" in the sense of Belletti–Rizzi (1988), Epstein et al. (1998), among others. The anaphor inside the subject is A-bound by the scrambled object in (36b).[24]

[24] É. Kiss (2002, ch. 3.4.2) discusses instrumental case marked arguments, like the ones in (i) below, arguing that grammatical functions as manifested in the form of case suffixes cannot be responsible for anaphora distribution, since—as she argues—there are examples (such as (i)) where it is the instrumental case-marked phrase that can bind the accusative anaphor, whereas in general it is the accusative argument that can bind the instrumental case-marked argument. However, if—as seems plausible—(i.a) is taken to have a structural description along the lines of (ii), a c-command based account of the distribution of anaphors in (i) is derived. Note that, although É. Kiss marks (i.b) as ungrammatical, it actually has a reading, where the pattern in (i.b) is well-formed: on that reading the

Scrambling also feeds pronominal variable binding, both in Hungarian (see (39a–b)) and in Japanese (see, e.g., Saito 2003, 485) (but not in German, see Grewendorf–Sabel 1999 (= G&S 1999)), to which the same explanation will extend.

(39) (a) *?EBBEN A VÁROSBAN bántalmazott [*pro* pár diákja]
this-in the town-in assaulted *pro*$_i$ several student-poss.3sg-nom
[kevés tanárt]
few teacher-acc$_i$

(b) EBBEN A VÁROSBAN bántalmazott [kevés tanárt]$_i$ [*pro*$_i$ pár diákja] t_i
'It's this town where few teachers were assaulted by several of their students.'

5.2. Scrambling and Condition C

Postverbal scrambling in Hungarian does not feed or obviate Condition C:

(40) (a) Látta (ön)magát$_i$ János t_i a tükörben
saw-3sg (his-)himself-acc$_i$ J.-nom t_i the mirror-in
'John saw himself in the mirror.'

(b) **Látták a fiúk anyját$_i$ ők t_i (cf. the discussion of (5))
saw-3pl the boy-pl$_i$ mother-acc they-nom$_i$ t_i
'*They$_i$ saw the boys'$_i$ mother.'

The same holds true of Japanese short (i.e., local) scrambling:

---

instrumental case-marked phrase is indeed the thematic instrument argument of the verb. Such an interpretation is illustrated in (iii).

(i) (a) A lányokkal felhívattam egymást.
the girls-with up-call-caus-past-1sg each other-acc
'I got the girls call each other.'

(b) *A lányokat felhívattam egymással.
the girls-acc up-call-caus-past-1sg each other-with
'*I got each other call the girls'

(ii) [I CAUSE [the girls-with call each other-acc]]

(iii) (a) Hívasd fel őket egymással!
call-caus-imp-2sg up them each other-with
'Make them call each other!'

(b) Kend meg a kenyereket egymással!
smear-imp pv the bread-pl-acc each other-with
'Butter the slices of bread with each other!'

(41) (a) [Zibunzisin-o$_i$ [John-ga t_i semeta]]
himself-acc$_i$ J.-nom t_i blamed
'John blamed himself.'

(b) *[[John-no hahaoya]-o$_i$ [kare-ga t_i semeta]]
J.-gen mother-acc$_i$ he-nom t_i blamed
'*He$_i$ blamed John's$_i$ mother.'

5.3. Scrambling and WCO

The Hungarian short scrambling operation does not induce WCO effects, rather, it obviates WCO violations. This was demonstrated by examples (23)–(24) in section **3.1** above. Note that if the object universal QP moves only as far as the t_i' position in (24) (object scrambling without the extra QP-fronting step in (24)), the result is still grammatical, see (42) below).

(42) (?)Felismerte ['mindegyik 'lányt]$_i$ [az a férfi, aki bement
recognized-3sg every girl-**acc**$_i$ that the man-**nom** who in-went-3sg
pro$_{3SG_i}$ hozzá] t_i
to.her$_i$
'*?The man who dropped by her$_i$ recognized every girl$_i$.'

Japanese type scrambling (and also German type scrambling, see (44)) exhibits analogous contrasts:[25]

(43) (a) ?*[[Soitu$_i$-no hahaoya]-ga [dare$_i$-o aisiteru]] no?
the-guy-gen mother-nom who-acc love Q
'?*Who$_i$ does his$_i$ mother love t_i?'

(b) ?Dare$_i$-o [[soitu$_i$-no hahaoya]-ga [t_i aisiteru]] no?
who-acc the-guy-gen mother-nom love Q (Saito 1992, 73)

(44) (a) *... weil seine$_i$ Mutter jeden Studenten$_i$ liebt
since his mother-nom every student-acc loves
'*His$_i$ mother loves every student$_i$.'

[25] Licensing of parasitic gaps (which is taken to be a property of A-bar movement) is notoriously difficult to test in Hungarian, but to the extent it is testable, it appears not to be affected by scrambling. If so, this would contrast Hungarian scrambling with German (and Dutch) scrambling, where parasitic gaps are apparently licensed by the scrambling movement. As for Japanese, parasitic gaps do not exist in the language (see Saito 1992).

(b) ... weil [jeden Studenten$_i$ [seine Mutter$_i$ t_i liebt]]
since every student-acc his mother$_{nom}$ loves (G&S 1999)

5.4. Scope

In cases where scope interpretation in a subject–object order is unambiguously S > O, as in (34a) above, reproduced here as (45), scrambling of the object over the subject introduces scope ambiguity, as in (46).

(45) TAVALY végzett el minden diák kevés kurzust
last.year did-3sg pv every student-nom few course-acc
'It was last year that every student did few courses.' (S > O, *O > S)

(46) TAVALY végzett el [kevesebb mint öt kurzust]$_i$ minden diák t_i
last.year did-3sg pv fewer than five course-**acc**$_i$ every student-**nom** t_i
'It was last year that every student did fewer than 5 courses.' (S > O, O > S)

The same holds true of Japanese, and German too. (45) illustrates the case for Japanese: while in the subject–object order only a direct scope interpretation is available, when the object is scrambled to the left of the subject, both scope interpretations become available.

(47) (a) Dareka-ga daremo-o aisite iru.
someone-nom everyone-acc loves
'Someone loves everyone.' $\exists > \forall$ / $*\forall > \exists$

(b) Daremo-o$_i$ dareka-ga t_i aisite iru
everyone-acc someone-nom loves
'Someone loves everyone.' $\exists > \forall$ / $\forall > \exists$

All in all, the basic properties of the postverbal reordering under scrutiny here appear to most closely match those of Japanese short scrambling.[26]

[26] Japanese short-scrambling is often categorized as A-scrambling (see Grewendorf–Sabel 1999 for corroboration of this view), but as it is well-known, at least prima facie, its properties are mixed (also involving traits of obligatory reconstruction, a putative property of (some) A-bar movements, which is uncharacteristic of A-movements; see Ueyama 2002 and Saito 2003 for two different approaches to this mixed behavior). Therefore I refrain here from situating Hungarian scrambling within the A/A-bar dichotomy (a distinction called into question in the current minimalist framework). My claim is simply that the basic properties of Hungarian scrambling, as reviewed in this section, reveal that the reordering operation behaves on a par with Japanese short scrambling.

This provides strong confirmation for the proposal that this reordering indeed involves scrambling in Hungarian.[27]

Finally, it is shown in Surányi (in press) that a number of fundamental implications involving scrambling that have been noted in the literature (e.g., V-raising, *pro*-drop, richness of morphology, etc.) are applicable to the language, which lends further plausibility to the scrambling approach I advocate.

6. Concluding remarks

The main result of the present paper is that it eliminates an alleged residual idiosyncrasy of Hungarian, the non-configurationality of its verb phrase, by demonstrating systematically that a scrambling approach, based on a configurational *v*P, is indubitably available, and what is more, empirically superior to the flat VP account. Modulo scrambling, Hungarian is configurational not only in its left periphery, but all the way down. It has also been shown that postverbal object–subject reordering in this language is akin in particular to short scrambling of the Japanese-type (and contrasts in crucial ways with German or Slavic scrambling). Due to verb raising, what occurs in the best-studied scrambling languages to the left of the verb characterizes the postverbal field in Hungarian.

It is not the purpose of this work to choose from, or evaluate, alternative approaches to Japanese/Hungarian-type local scrambling. At

[27] It is not clear if a specific interpretation of indefinites should be enforced in a scrambled position (as in Dutch or German, see de Hoop 1992) (e.g., examples like *Keres egy ügyvédet Mária* 'lit. seeks a lawyer-acc Mary-nom' appear to be degraded for some speakers if the indefinite object NP is non-specific, but judgments are murkier in other cases.) Nevertheless, options for the projection of information focus are affected by scrambling in much the same way as in Japanese (cf. Miyagawa 2005 and references therein, see also Neeleman–Reinhart 1998 for a discussion of Dutch). A sentence like (i) can answer either *What happened?* or *What did John do?* or *Who did John see?*, whereas (ii) involving the scrambled order is apparently inappropriate as an answer to the last two questions. (ii) can serve as a (non-exhaustive) answer to *Who saw the teacher?*, whereas (i) is not felicitous in the same context.

(i) Meglátta János a tanárt
pv-saw John-nom the teacher-acc

(ii) Meglátta a tanárt János
pv-saw the teacher-acc John-nom

the same time, insofar as the main conclusions reached here can be upheld, a more microscopic study of the properties of Hungarian scrambling can serve as excellent testing ground for current competing accounts of Japanese-type scrambling, with repercussions for the ongoing debate over the proper typology of scrambling in general.

References

Ariel, Mira 1994. Interpreting anaphoric expressions: A cognitive versus a pragmatic approach. In: Journal of Linguistics 30 : 1–42.

Baker, Mark C. 1988. Incorporation: A theory of grammatical function changing. University of Chicago Press, Chicago.

Barbosa, Pilar – Danny Fox – Paul Hagstrom – Martha McGinnis – David Pesetsky (eds) 1998. Is the best good enough? Optimality and competition in syntax. MIT Press, Cambridge MA.

Belletti, Adriana – Luigi Rizzi 1988. Psych-verbs and θ-theory. In: Natural Langauge and Linguistic Theory 6 : 291–352.

Bennis, Hans 1986. Gaps and dummies. Foris, Dordrecht.

Boškovič, Želko 2002. On multiple *wh*-fronting. In: Linguistic Inquiry 33 : 351–84.

Boškovič, Želko – Daiko Takahashi 1998. Scrambling and last resort. In: Linguistic Inquiry 29 : 347–66.

Bresnan, Joan 1995. Linear order, syntactic rank, and empty categories: On weak crossover. In: Mary Dalrymple – Ronald M. Kaplan – John T. Maxwell – Annie Zaenen (eds): Formal issues in Lexical-Functional Grammar, 241–74. Stanford University, CSLI Publications.

Bresnan, Joan 1998. Morphology competes with syntax: Explaining typological variation in Weak Crossover effects. In: Barbosa et al. (1998, 59–92).

Brody, Michael 1990. Remarks on the order of elements in the Hungarian focus field. In: István Kenesei (ed.): Approaches to Hungarian 3 (Structures and arguments), 95–121. JATE, Szeged.

Brody, Michael 1995. Focus and Checking Theory. In: István Kenesei (ed.): Approaches to Hungarian 5 (Levels and structures), 29–43. JATEPress, Szeged.

Brody, Michael – Anna Szabolcsi 2003. Overt Scope in Hungarian. In: Syntax 6 : 19–51.

Choe, Hyon-Sook 1989. Restructuring parameters and scrambling in Korean and Hungarian. In: László Marácz – Peter Muysken (eds): Configurationality: The typology of asymmetries, 267–92. Foris, Dordrecht.

Chomsky, Noam 1993. A Minimalist Program for Linguistic Theory. In: Hale – Keyser (1993a, 1–52).

Chomsky, Noam 1995. The minimalist program. MIT Press, Cambridge MA.

Chtareva, Angelina 2005. An experiencer analysis of subject idioms in Russian. In: Proceedings of FASL 13. The Columbia Meeting, 80–91. Michigan Slavic Publications, Ann Arbor.

Dobrovie-Sorin, Carmen 1990. Clitic doubling, *wh*-movement, and quantification in Romanian. In: Linguistic Inquiry 21 : 351–97.

É. Kiss, Katalin 1987a. Configurationality in Hungarian. Akadémiai Kiadó, Budapest.

É. Kiss, Katalin 1987b. Is the VP universal? In: István Kenesei (ed.): Approaches to Hungarian 2 (Theories and analyses), 13–85. JATE, Szeged.

É. Kiss, Katalin 1991. Logical structure in syntactic structure: The case of Hungarian. In: James Huang – Robert May (eds): Logical structure and linguistic structure: Cross-linguistic perspectives, 111–48. Kluwer, Dordrecht.

É. Kiss, Katalin 1994a. Scrambling as the base-generation of random complement order. In: Norbert Corver – Henk van Riemsdijk (eds): Studies on scrambling, 221–56. Mouton de Gruyter, Berlin & New York.

É. Kiss, Katalin 1994b. Sentence structure and word order. In: Katalin É. Kiss – Ferenc Kiefer (eds): The syntactic structure of Hungarian (Syntax and semantics 27), 1–90. Academic Press, New York.

É. Kiss, Katalin 1995. NP-movement, WH-movement and scrambling in Hungarian. In: Katalin É. Kiss (ed.): Discourse configurational languages, 207–43. Oxford University Press, Oxford.

É. Kiss, Katalin 2002. The syntax of Hungarian. Cambridge University Press, Cambridge.

É. Kiss, Katalin 2003. Argument scrambling, operator movement and topic movement in Hungarian. In: Karimi (2003, 22–43).

É. Kiss, Katalin 2006. A phase-based theory of free word order. Manuscript. Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest.

Epstein, Samuel David – Erich M. Groat – Ruriko Kawashima – Hisatsugu Kitahara 1998. A derivational approach to syntactic relations. Oxford University Press, Oxford.

Everaert, Martin 1993. Verbal idioms, subject idioms and theta-theory. In: Laurel Smith Stvan (ed.): Papers From the Third Annual Formal Linguistics Society of Midamerica Conference, 43–59. IULCP, Bloomington.

Fanselow, Gisbert 2001. Features, theta-roles, and free constituent order. In: Linguistic Inquiry 32 : 405–37.

Fanselow, Gisbert 2004. The MLC and derivational economy. In: Arthur Stepanov – Gisbert Fanselow – Ralf Vogel (eds): The minimal link condition, 74–123. Mouton de Gruyter, Berlin & New York.

Fox, Danny 1995. Economy and scope. In: Natural Language Semantics 3 : 283–341.

Fox, Danny 1998. Locality in variable binding. In: Barbosa et al. (1998, 129–55).

Fox, Danny 2000. Economy and semantic interpretation. MIT Press, Cambridge MA.

Fukui, Naoki 1993. Parameter and optionality in a grammar. In: Linguistic Inquiry 24 : 399–420.

Grewendorf, Günther – Joachim Sabel 1999. Scrambling in German and Japanese: Adjunction versus multiple specifiers. In: Natural Language and Linguistic Theory 17 : 1–65.

Grohmann, Kleanthes K. 1997. German superiority. In: Abraham Werner – Kleanthes K. Grohmann (eds): Groninger Arbeiten zur germanistischen Linguistik 40, 97–107. University of Groningen, Groningen.

Haider, Hubert 1986. Deutsche Syntax – Generativ. Ph.D. thesis, Habilitation thesis, Vienna.

Hale, Kenneth – Samuel J. Keyser (eds) 1993a. The view from building 20. Essays in linguistics in honor of Sylvain Bromberger. MIT Press, Cambridge MA.

Hale, Kenneth – Samuel Jay Keyser 1993b. On argument structure and the lexical expression of syntactic relations. In: Hale – Keyser (1993a, 53–109).

Heycock, Caroline 1995. Asymmetries in reconstruction. In: Linguistic Inquiry 26: 547–70.

de Hoop, Helen 1992. Case configuration and NP interpretation. Doctoral dissertation, University of Groningen.

Horváth, Júlia 1986. Focus in the theory of grammar and the structure of Hungarian. Foris, Dordrecht.

Ishii, Toru 2006. A nonuniform analysis of overt *wh*-movement. In: Linguistic Inquiry 37: 155–67.

Karimi, Simin (ed.) 2003. Word order and scrambling. Blackwell, Malden MA & Oxford.

Kayne, Richard 1994. The antisymmetry of syntax. MIT Press, Cambridge MA.

Kenesei, István 1984. On what really figures in a non configurational language. In: Groninger Arbeiten zur germanistischen Linguistik 24: 28–54.

Kenesei, István 1986. On the logic of Hungarian word order. In: Werner Abraham – Sjaak de Meij (eds): Topic, focus, and configurationality, 143–59. John Benjamins, Amsterdam.

Kenesei, István 1989. On pronominal binding in Hungarian. In: László Marácz – Peter Muysken (eds): Configurationality: The typology of asymmetries, 223–36. Foris, Dordrecht.

Kitahara, Hisatsugu 2002. Scrambling, case, and interpretability. In: Samuel David Epstein – Daniel T. Seely (eds): Derivation and explanation in the Minimalist Program, 167–79. Blackwell, Malden MA & Oxford.

Ko, Heejeong 2005. Syntax of why-in-situ: Merge into [Spec,CP] in the overt syntax. In: Natural Language and Linguistic Theory 23: 867–916.

Komlósy, András 1983. Vonzatok és szabad határozók az egyszerű mondatokban [Complements and adjuncts in simple sentences]. Manuscript, Research Institute for Linguistics, Hungarian Academy of Sciences.

Koopman, Hilda – Dominique Sportiche 1982. Variables and the Bijection Principle. In: The Linguistic Review 2: 139–60.

Kuno, Susumu 1982. The focus of the question and the focus of the answer. In: CLS 18: Papers from the Parasession on Nondeclaratives, 134–57. Chicago Linguistics Society, Chicago.

Lasnik, Howard – Tim Stowell 1991. Weakest crossover. In: Linguistic Inquiry 22: 687–720.

Marantz, Alec 1984. On the nature of grammatical relations. MIT Press, Cambridge MA.

Marantz, Alec 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In: UPenn Working Papers in Linguistics 4.2: 201–25.

Marácz, László 1989. Asymmetries in Hungarian. Doctoral dissertation, Groningen University.

McGinnis, Martha 2004. Lethal ambiguity. In: Linguistic Inquiry 35:47–95.

Miyagawa, Shigeru 1997. Against optional scrambling. In: Linguistic Inquiry 28:1–25.

Miyagawa, Shigeru 2001. EPP, scrambling, and *wh*-in-situ. In: Michael Kenstowicz (ed.): Ken Hale: A life in language, 239–338. MIT Press, Cambridge MA.

Miyagawa, Shigeru 2003. A-movement scrambling and options without optionality. In: Karimi (2003, 177–200).

Miyagawa, Shigeru 2005. On the EPP. In: Perspectives on phases. MIT Working Papers in Linguistics 49:201–35.

Neeleman, Ad – Tanya Reinhart 1998. Scrambling and the PF interface. In: Miriam Butt – Wilhelm Geuder (eds): The projection of arguments lexical and compositional factors, 309–53. CSLI Publications, Stanford.

Nunberg, Geoffrey – Ivan Sag – Terry Wasow 1994. Idioms. In: Language 70:491–538.

Perlmutter, David 1971. Deep and surface structure constraints in syntax. Holt, Rinehart and Winston, New York.

Postal, Paul 2002. Skeptical linguistic essays. Manuscript. NYU.

Puskás, Genoveva 2000. Word order in Hungarian. The syntax of A-bar positions. John Benjamins, Amsterdam & Philadelphia.

Reinhart, Tanya 1983. Anaphora and semantic interpretation. Croom Helm, London.

Reinhart, Tanya 1995. Interface strategies (OTS Working Papers in Theoretical Linguistics). OTS, Utrecht.

Reinhart, Tanya 1998. Interface economy: Focus and markedness. In: Chris Wilder – Hans-Martin Gaertner – Manfred Bierwisch (eds): The role of economy principles in linguistic theory, 146–69. Akademie-Verlag, Berlin.

Reinhart, Tanya 2006. Interface strategies. Reference set computation. MIT Press, Cambridge MA.

Richards, Norvin 1997. What moves where in which language? Doctoral dissertation, MIT.

Rizzi, Luigi 2001. Reconstruction, weak island sensitivity and agreement. In: Carlo Ceccheto – Gennaro Chierchia – Maria Teresa Guasti (eds): Semantic interfaces, 145–76. CSLI, Stanford.

Rudin, Catherine 1988. On multiple questions and multiple *wh*-fronting. In: Natural Language and Linguistic Theory 6:445–501.

Ruys, Eddy 1994. A global economy account of Weak Crossover. In: Linguistics in the Netherlands 1994:223–34.

Sabel, Joachim 2002. A minimalist analysis of syntactic islands. In: The Linguistic Review 19:271–315.

Saito, Mamoru 1985. Some asymmetries in Japanese and their theoretical implications. Doctoral dissertation, MIT.

Saito, Mamoru 1992. Long distance scrambling in Japanese. In: Journal of East Asian Linguistics 1:69–118.

Saito, Mamoru 2003. A derivational approach to the interpretation of scrambling chains. In: Lingua 113:481–518.

Saito, Mamoru – Naoki Fukui 1998. Order in phrase structure and movement. In: Linguistic Inquiry 29:439–74.

Shavrit, Yael – Elena Guerzoni 2003. Reconstruction and its problems. In: Paul Dekker – Robert van Rooy (eds): Proceedings of the 14th Amsterdam Colloquium, 205–10. University of Amsterdam, Amsterdam.

Speas, Margaret J. 1990. Phrase structure in natural language. Kluwer, Dordrecht.

Spector, Benjamin 2004. Pseudo weak crossover in French relative clauses and global economy. In: Actes des Journées d'Etudes Linguistiques de Nantes (JEL04), 155–64. University of Nantes, Nantes.

Surányi, Balázs 2002. Multiple operator movements in Hungarian. Doctoral dissertation, LOT, Utrecht.

Surányi, Balázs 2004. Differential quantifier scope: Q-raising versus Q-feature checking. In: Empirical Issues in Syntax and Semantics 5:215–40.

Surányi, Balázs 2006. Mechanisms of *wh*-saturation and interpretation in multiple *wh*-movement. In: Lisa Cheng – Johan Rooryck (eds): WH-movement moving on, 289–318. MIT Press, Cambridge MA.

Surányi, Balázs in press. Hungarian as a Japanese-type scrambling language. In: Christopher Davis – Amy Rose Deal – Youri Zabbal (eds): Proceedings of NELS 36. GLSA Publications, University of Massachusetts, Amherst.

Szabolcsi, Anna 1983. The possessor that ran away from home. In: The Linguistic Review 3:216–89.

Szabolcsi, Anna 1997. Strategies for scope taking. In: Anna Szabolcsi (ed.): Ways of scope taking (SLAP 65), 109–55. Kluwer, Dordrecht.

Szczegielniak, Adam 1999. That-t effects cross-linguistically and successive cyclic movement. In: MIT Working Papers in Linguistic 33:369–93.

Ueyama, Ayumi 2002. Two types of scrambling constructions in Japanese. In: Andrew Barss (ed.): Anaphora: A reference guide, 23–71. Blackwell, Malden MA & Oxford.

Varga, László 1981. A topicról és a fókusz utáni elemek sorrendjéről [On the topic and the order of elements after the focus]. In: Magyar Nyelv 27:198–200.

Wiltschko, Martina 1998. Superiority in German. In: Emily Curtis – James Lyle – Gabriel Webster (eds): Proceedings of WCCFL 16, 431–46. CSLI, Stanford.

Acta Linguistica Hungarica, Vol. 53 (4), pp. 433–447 (2006)
DOI: 10.1556/ALing.53.2006.4.3

THE MARKEDNESS OF THE UNMARKED*

PÉTER SZIGETVÁRI

Research Institute for Linguistics
Hungarian Academy of Sciences
Benczúr utca 33.
H–1068 Budapest
Hungary
peter.szigetvari@elte.hu

Abstract: There are several phonological categories whose markedness—as inferred from typical markedness metrics—fails to match the representational complexity posited for them. More specifically, glottal stops, geminate clusters, and onsetless syllables are representationally the simplest of their category, yet other criteria, like implicational hierarchies, mark them as special. This paper aims at comprehending this paradox.

Keywords: markedness, complexity, implicational hierarchies, place of articulation, consonant cluster, syllable structure

1. Introduction

Markedness is a hot issue in phonological theory, one barely dares touch it. The reason for much of the excitement surrounding the concept of markedness is its usefulness: it is very appealing to base theory-internal considerations—like, for example, the representation of segments, or clusters—on what looks like theory-external evidence—like markedness. The greatest difficulty with the concept is the lack of a consensus on what counts and what does not count for determining whether a given cate-

* I received very useful comments from Péter Siptár, for which I am grateful. I was supported by the Hajdú Péter grant of the Research Institute of Linguistics, Hungarian Academy of Sciences.

1216–8076/$ 20.00

gory is marked or unmarked. This leads Hume, for example, to claim that "markedness need not be encoded in the grammar" (2003, 16).

Rice (2003; to appear) collects a long list of properties that phonologists use to characterize marked vs. unmarked categories. These include a set she refers to as nonphonological properties: e.g., language-internal and cross-linguistic **frequency** (rare categories are marked, common ones unmarked), **psycholinguistic** observations (categories acquired early and lost late are unmarked), **implicational** relations (a marked category implies the presence of its unmarked counterpart in the system), difficulty of **articulation** and more **perceptual** salience of marked categories as compared to unmarked ones. The second set of criteria is labelled phonological by Rice, this includes **neutralization,** of which marked categories are typically the target, unmarked ones the result, **epenthesis,** which typically involves unmarked segments to the exclusion of marked ones, **assimilation,** of which marked categories are the trigger and unmarked ones the target, as well as **coalescence** and **deletion,** in which marked categories are retained, unmarked ones lost.

In an ideal world, the markedness relations that are established on the basis of the empirical criteria available sketch up a uniform pattern, i.e., a given category always emerges as more marked than another category, irrespective of which of the above criteria are applied. Furthermore, if our theory dictates that a given category is more marked than another category, it would be nice to see that the discriminatory properties listed above support the theory's statement about markedness.

Our world, however, is not ideal. It happens all too often that a category is deemed unmarked by one criterion, but marked by another. To mention some well-known examples: the glottal stop, [ʔ], is a common result of the neutralization of oral plosives, hence allegedly unmarked. Yet its presence in a system almost always implies the presence of the coronal plosive, [t] — while the opposite implication does not hold —, hence the glottal stop is supposed to be more marked than the coronal one. A syllable containing only a vowel evidently contains less marks, only the properties of the vowel, than one beginning with a consonant, which also contains the properties of the consonant. It is nevertheless the latter that acquired the reputation of unmarked syllable type, since it does not imply the onsetless syllable, while languages with onsetless syllables exceptionlessly have onsetful syllables as well.

I will have very little to say about most of the contradictory verdicts of markedness criteria. My aim is only to show that in a number of

cases what appears to be the least marked category fails to take a role in a linguistic system. Thus I try to interpret the observation that some categories which fall out as unmarked in one sense, behave as marked in the other (e.g., Wilson 2001, 151), a general excuse for which is that markedness is multidimensional. I will argue instead that two concepts, markedness and complexity, are merged in such discussions. If my claim holds, a single scale of markedness can be retained. This scale coincides with that of complexity to a very large extent, but the two **may** depart at the lower end: the least complex category may or may not be equivalent to the least marked one in a given system.

I will first present three paradoxes that can be brought into parallel. Each demonstrate that minimal complexity is marked. The next section aims at clarifying the relationship of markedness and complexity. Section 4 discusses a possible complexity metric compatible with the notion of markedness and complexity presented here.

2. The problem

In the following three sections I will introduce three paradoxes that look stunning when considered from the viewpoint of markedness. The first involves the composition of segment inventories, the second that of consonant clusters. The third case offers a very similar example from the domain of syllabic constituency. All three cases demonstrate that the category at the lower end of what looks like a well-founded complexity (or markedness) scale fails to turn up in a large set of languages. This is odd if we expect unmarked categories to be very common in linguistic systems.

2.1. Segments

Based on the UPSID corpus (Maddieson 1984), Rebrus and Trón (2004) show that all of the 317 languages have at least three released plosives, i.e., plosives come in at least three different places of articulation without exception in a representative sample of human languages. A further generalization can be made about these three places of articulation: they are coronal, velar, and labial ([t k p]). There are only five languages (1.6% of the sample) in which we do not find all three members of this canonical triplet; one of them is missing in four—coronal in Hawaiian,

velar in Kirghiz, and labial in Aleut and Wichita—and two, both velar and labial, in Hupa. The deficiency is made up for by a glottal, a palatal, a uvular, a labialized velar stop, or some combination of these, since, recall, there are at least three in all the languages in the database.

The chart in (1)—adopted from Rebrus and Trón's (3)—summarizes the above statements. The authors claim that the chart also represents a complexity scale, where explosives to the right are more complex (that is, marked) than those to the left. Arrows indicate the range of places that feature in the given language. An asterisk marks the absence of that place, a tilde marks variability: the relevant category may or may not turn up in the language. The languages are listed exhaustively relative to the UPSID corpus.

(1)

?	t/d	k/g	p/b	others	languages
◁——	*	——▷			Hawaiian
	◁——	*	——▷		Kirghiz
◁——			*	——▷	Wichita
	◁——		*	——▷	Aleut
◁——		*	*	——▷	Hupa
~	◁——		——▷	~	312 others

Rebrus and Trón argue that by merging the three common places of articulation (coronal, velar, and labial) we get a complexity scale on which the principle in (2) holds. Without the merger, the five languages mentioned in (1) will exhibit exceptional inventories.

(2) The phoneme inventory of all languages is convex (not discontinuous) on any complexity scale.

Put somewhat differently:

(3) The phoneme inventory of a given language is unambiguosly defined by the elements of minimal and maximal complexity on any complexity scale.

Were it not for the fact that the presence of the glottal stop in an inventory cannot be predicted, a much stronger claim could be made, this is given in (4).

(4) The phoneme inventory of a given language is unambiguosly defined by the element of maximal complexity on a given complexity scale.

Such a principle is equivalent to the implicational hierarchies mentioned above: some categories imply the presence of others in a system, more

precisely, more complex—or more marked—categories imply the presence of less complex—or less marked—categories. E.g., [c/q] ⊃ [t/k/p] ⊃ [ʔ]. The data in (1), however, show that the stronger principle in (4), and hence the implicational chain of plosive places of articulation, cannot be maintained, since [t/k/p] ⊅ [ʔ], in fact, [ʔ] ⊃ [t/k/p].

The reverse implicational relationship of the glottal stop and the canonical plosives hints at the need to swap the two categories. Other considerations, however, militate against this: consonant lenition very often targets oral plosives and results in a glottal stop; consonant epenthesis also frequently prefers the glottal to the other places of articulation.

By surveying cases of consonant epenthesis, Lombardi (2002) aims to tackle the problem that despite the fixed place markedness hierarchy she proposes—similar to Rebrus and Trón's, shown in (5)—, a coronal may occur as an epenthetic consonant.[1]

(5) *dorsal, *labial ≫ *coronal ≫ *pharyngeal

To save the hierarchy, Lombardi argues that most cases of coronal epenthesis are not genuine, either because they are motivated by morphological factors, or because there exist further constraints on the properties of the epenthetic consonant (e.g., it must be a sonorant), so that pharyngeals do not qualify (as hinted at above: laryngeals are a subset of pharyngeals, and both [ʔ] and [h] are obstruents in Lombardi's view). Yet there remain cases where she has to assume a further constraint specifically against the glottal stop, which is somewhat simplistically labelled as *ʔ. About this constraint, Lombardi admits that it "is obviously contradictory to the unmarkedness of the major Place of the glottal stop" (*op.cit.*, 239).

With the aim of discrediting the notion of markedness altogether, Hume (2003) shows that besides the usual laryngeal, coronal, and velar places of articulation, the less usual labial may also be diagnosed as unmarked. Thus all the three canonical places, as well as the glottal stop exhibit symptoms characteristic of unmarkedness. Crucially, no such evidence was provided so far for other places (like palatal, uvular, etc.).[2]

[1] ≫ means 'dominates', that is, ranks higher in the constraint hierarchy. Accordingly, it is universally worse to have a dorsal or labial consonant than to have a coronal one, and it is universally worse to have a coronal than a pharyngeal (in our case laryngeal) consonant—Lombardi stipulates a ranking *[−glottal] ≫ *[+glottal] to achieve laryngeal unmarkedness (2002, 222).

[2] If Lombardi's (2003) generalizations about epenthetic vowels hold, a very similar situation can be witnessed for vowels. The claim is that the optimal epenthetic

In order to retain the notion of markedness in phonological theory, we must face the question: why do not all languages have a glottal stop in their phonemic inventory?

2.2. Clusters

Consonant clusters do not constitute a uniform set. A major, albeit fuzzy, split across the category is defined by the sonority profile of the cluster. Clusters with a rising or level sonority profile imply clusters with a falling sonority profile (Kaye–Lowenstamm 1981, 291,[3] also cf. Charette 1992 and Cyran 2003) Let us concentrate only on the unmarked set, and within this set only clusters with a plosive in second position. Furthermore, the present discussion must limit itself to intervocalic clusters. It is well-known that clusters are context sensitive: different types occur word initially and word finally. Both initial and final clusters—and potentially even others—occur word medially between vowels.

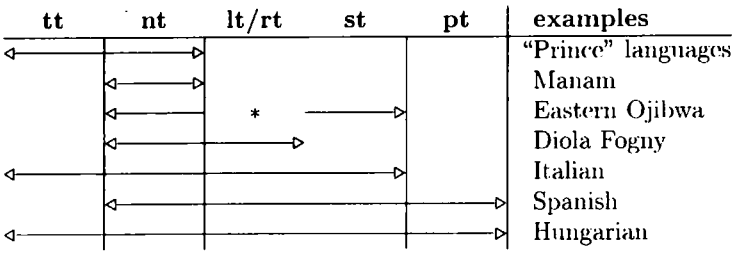
So-called "Prince" languages (Prince 1984; Harris 1997), possess a very modest set of consonant clusters: geminates and homorganic nasal + plosive clusters. Prince mentions Southern Paiute and Japanese as belonging to this group (1984, 243). While geminates do not, homorganic nasal + plosive clusters occur in any language that allows consonants to cluster. We may conclude then that this is the least marked type of consonant cluster in intervocalic position.

In fact, types of consonant cluster can be organized on a complexity (or markedness) scale in much the same way as individual segments, that is, one can produce scales based on various considerations which others can then dispute. (6) displays an adaptation of Rebrus and Trón's similar chart, omitting some details deemed unnecessary for the current discussion.

vowels are [ɨ] and [ə], the canonical triplet [i/a/u] are epenthesized only if the former are absent in the system. Vowel inventories, on the other hand, are based on the canonical three vowels, with the central ones occurring only as extensions. Thus the unmarked central vowels are missing from many vowel inventories, just like glottal consonants from consonant inventories.

[3] Kaye and Lowenstamm claim that there is an implicational relationship between the branchingness of syllabic constituents: onsets and nuclei branch only in languages in which the rhyme branches, that is, the least marked consonant cluster type is the product of a coda (the right branch of a branching rhyme) and the following onset consonant. Such clusters typically exhibit a falling sonority profile.

(6)



The clusters in the chart exemplify geminates ([tt]), homorganic nasal + plosive clusters ([nt]), homorganic liquid + plosive clusters ([lt/rt]), sibilant + coronal plosive clusters ([st]), and nonhomorganic clusters ([pt]). The chart abstracts away from a number of details, like voicing or place of articulation. Some languages allow only voiceless geminates (Kirchner 2000); in nasal + plosive clusters a voiced plosive is less marked, while in liquid + plosive, and especially obstruent + plosive clusters voiceless plosives occur more frequently. Place of articulation also lends some excitement to such a scale, but considering the simplest cases is enough for our present purposes.

To maintain the convexity of cluster inventories, Rebrus and Trón claim that the two categories liquid + plosive and *s* + plosive have to be merged, similarly to the merger executed in (1). The consequence of claiming that liquid + plosive and *s* + plosive clusters are equally complex is that neither implies the other in a given system, which is in fact the case: Eastern Ojibwa has *s* + plosive but no liquid + plosive clusters, while Diola Fogny exhibits the opposite setting.

As in the case of individual segments, where what is claimed to be the most unmarked stop, the glottal stop is not implied by any other place of articulation, here too geminates occur or not irrespective of the other clusters of the langauge. If markedness were based simply on implications, geminates could not count as unmarked, since they are not implied by other clusters: hosts of well-known languages with the most complex types of consonant clusters imaginable lack true geminates. Yet, the representation of geminates given in (11) below, which enjoys wide-spread popularity, hints at this type of cluster being the simplest possible. We again face the question: why do not all languages that have consonant clusters in the first place have true geminates?

2.3. Syllabic constituents

The unmarked syllable type cross-linguistically is CV. There exist languages (e.g., Hua (Blevins 1995, 219)) which only possess this type, that is, all words are of the shape $(CV)+$ (where '+' means one or more occurrences of the preceding pattern).

Although Blevins says "the unmarked case is that onsets are not obligatory" (*op.cit.*, 220), her own criteria rebut this claim. She brings up four arguments to support statements about unmarkedness in syllabic constituency. Of these three explicitly argue **for** obligatory onsets being the default case: "(3) All languages have CV syllables" — while, apparenly, only **some** have V syllables. Accordingly, CV is less marked than V. If it were the marked case to have obligatory onsets, then languages with only CV syllables would be more marked than others with both CV and V syllables. The oddity is that while the former group has only the unmarked syllable type (CV), the latter has also a marked type (V). Also, "(4) [...] there are a variety of phonological processes which take marked syllable types to unmarked types [...], but there are few if any rules which consistently result in [marked syllable types]." The avoidance of hiatus is a widespread phenomenon, which aims at getting rid of onsetless syllables (cf. Siptár 2006), while losing onsets is much less typical — though not unprecedented — a process. Blevins also says: "(2) In second language acquisition, speakers have little difficulty in shifting from a 'yes' value to a 'no' value for a given parameter, but do show difficulty in switching from a 'no' value to a 'yes' value." Yet it is hard to see any difficulty in getting from a language having both CV- and V-type syllables to one which only has the former, which allegedly is the marked type. Furthermore, the "yes" and "no" values crucially depend on the formulation of the parameter: "obligatory onset" for Blevins, hence her feeling that "no" is the unmarked setting, but it might as well be "omissible onset", in which case it is indeed the "no" setting that is unmarked.

This means that in the default case the onset in an obligatory part of the syllable (e.g., Prince–Smolensky 1993, 89). Onsets may contain more than one consonant, e.g., [tr] in English *trap* or Italian *tra* 'between'. Such more complex versions of onset are again restricted to a subset of human languages, hence count as marked.

Looking at codas on the other hand, we find that codaless syllables are omnipresent in human languages. This asymmetry clearly manifests itself in Clement's Length Hierarchy, given in (7).

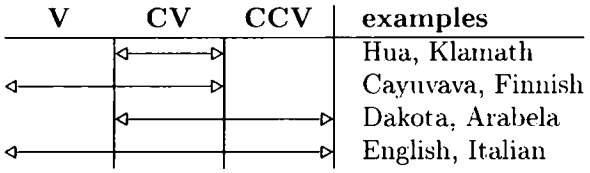
(7) The Length Hierarchy (Clements 1990, 307)

For any given type t, the presence in L(IS) of a demisyllable of length l ($l > 2$) implies the presence of a demisyllable of length $l - 1$.

According to the hierarchy, if a language after initial syllabification, L(IS), has a CCV initial demisyllable, it will also have a CV demisyllable, i.e., CCV $\supset$ CV. However, Clements has to add the constraint $l > 2$, since CV $\not\supset$ V: a two-long initial demisyllable does not imply the one-long initial demisyllable. For final demisyllables no such constraint is necessary, VCC $\supset$ VC $\supset$ V.[4] But this would require that there be two separate length hierarchies, one for initial and another for final demisyllables, the constraint holding only in the former. Instead, to gear his algorithm to linguistic facts, Clements stipulates that final V demisyllables are the least marked, while initial V demisyllables are more marked than any initial CV demisyllable. The raggedness of the theory is caused by the fact that the least complex onset type is not the least marked one.

To make the parallelism of this and the previous two cases obvious, the usual chart is sketched up in (8). The reason for the orderliness of this chart as compared to those in (1) and (6) above is simply that the identity of the consonants is ignored. If, say, the manner of the consonants were considered the result would be rather similar to the complicatedness of the other charts.

(8)



The example languages in the chart in (8) are from Blevins (1995, 219). Note that decisions on what a branching onset is are theory specific: some researchers consider any word-initial consonant cluster a branching onset, others would disagree with that. Irrespective of the actual examples (of whether Dakota and Arabela, or in fact any language with obligatory onsets indeed have branching onsets), the point holds: again the first column of the chart is independent of the others: the least complex type of onset is not implied by more complex onset types.

[4] The fact that codas are so easily dismissable has led to dispensing with it as a theoretical category in, for example, government phonology (Kaye et al. 1990).

Adherents of theories where not only onsets but also nuclei can be nil will notice that the same case can be made with respect to nuclei: the least complex nucleus, the empty one, is far from being unmarked. Thus, while a branching nucleus implies a nonbranching one—which is present in all languages—, the possibility of empty nuclei is not implied by either type of nucleus.

The paradox is well illustrated by the chart Kaye–Lowenstamm (1981, 292) produce for calculating the markedness of syllabic constituents. It is reproduced here in (9).

(9)

onset	rhyme	markedness
C	V	0
$\emptyset$	$\emptyset$	1
CC	VC	2
CCC	VCC	3
$C_1 \ldots C_n$	$VC_1 \ldots C_{n-1}$	n

While generally the more consonants in the onset the more marked the onset is, the lapse is clearly visible at the beginning of the scale: zero consonant is more marked than one consonant.

* * *

To summarize: we have seen three independent cases in which categories placed on a complexity scale defy the expectation that the least complex one (glottal stop, geminate consonant, onsetless syllable) be the least marked one as well. We have also seen that some of these facts have caused problems for researchers: the quest for the unmarked place of articulation resulted in four candidates—laryngeal, coronal, velar, and labial—, which made Hume, for example, conclude that markedness should be ousted of linguistic discussion.

3. What is marked?

The notion of markedness in linguistic theory goes back to Trubetzkoy,[5] who uses the term **mark** basically to mean distinctive feature (discriminative elements—in Baltaxe's translation). The four marks characteristic of *k*, for example, are "(1) complete occlusion" (noncontinuant), "(2) blocking of the entrance to the nasal cavity" (nonnasal), "(3) tightening of the

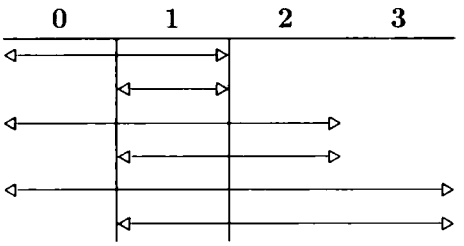
[5] Trubetzkoy attributes the term "mark" to Karl Bühler.

muscles of the tongue and simultaneous relaxation of the muscles of the larynx" (fortis), and "(4) participation of the dorsum" (dorsal) (1969, 66). Losing a mark will neutralize the difference between [k] and some other (group of) sound(s). For Trubetzkoy marks are language specific, since phonological contrasts vary from language to language. In contemporary models of phonological representation, however, the set of marks (also known as features, components, elements, gestures) is universal, and as a result the statements about markedness relations must also be universal.

While phonological models applying binary features are capable of encoding markedness considerations—as Chomky–Halle (1968, 402ff) and especially Kean (1975), as well as theories of underspecification show—a much more trivial interpretation of mark is available in models using privative features. In such a theory, each feature is a mark, accordingly the complexity (number of features) of a segment can be equated with its markedness.

The charts in (1), (6), and (8) can be merged as in (10).

(10)



The numbers stand for the complexity index of the given categories. Categories with complexity index 1 (including the canonical places of articulation, nasal + plosive clusters, CV syllables, etc.) are expected to turn up in all systems. As regards frequency and implicational relations, these are the most unmarked categories of the system. Categories with higher complexity indexes always imply categories with lower complexity indexes. Categories with complexity index 0 (glottals, geminates, onsetless syllables), on the other hand, are not implied by other categories, but they imply cateogries with complexity index 1. Thus many markedness diagnostics do not select them as unmarked, in fact, implicational hierarchies predict categories of complexity 0 to be more marked than categories on complexity 1. Yet if we consider their representation, i.e., their complexity, they do appear to be totally unmarked.[6]

[6] As a somewhat frivolous parallel, one can think of the distribution of matter in our universe. By far the most widespread form is hydrogen, which has an atomic

The following quote from Lombardi offers some explanation for the reason why zero complexity is discouraged in language:

> "I assume that this constraint [her aforementioned *?] is due to the perceptual difficulty of glottal stop [...]. The conflict between the formal unmarkedness of the Place of the glottal and its perceptual markedness accounts for the fact that while glottal stop is relatively unmarked, not all languages have it in their inventory, since they may resolve this conflict in different ways." (2002, 239)

We can conclude that (at least some of) the confusion surrounding the notion of markedness is terminological. Empirical diagnostics of markedness usually single out categories of complexity 1, but occasionally they point to categories of complexity 0. Complexity is a theoretical notion, the complexity of a category is theory dependent. Markedness, on the other hand, is an empirical issue. Theories can be assessed by the degree of accordance between the results of markedness diagnostics and the complexity values the theory posits.

4. Representation

The complexity metric for syllabic constituents is self-evident: it is not difficult to accept that the complexity of a (two-way) branching onset is 2, that of a nonbranching onset is 1, while that of a missing onset is 0.

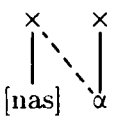
In the case of clusters, measuring complexity is less trivial. It is generally accepted that in the representation of a true geminate consonant it is to the second position that all melodic material (call them features) are anchored, it is this consonantal position that licenses features (Goldsmith 1990, 128ff). The first position parasitically links to the features held by the second, but does not license any on its own. This is indicated by the dashed association line in (11), where α stands for any (bundle of) features defining some segment.

(11)

number (let us say complexity) of 1. The second most common element is helium, with a complexity of 2. An element of complexity 0 can also be hypothesized, but is it so rare in our universe that there is only some faint empirical evidence of its existence (http://en.wikipedia.org/wiki/Tetraneutron).

To produce zero complexity for this structure, we must consider the number of features anchored to the first position independently of the second.

In a homorganic nasal + plosive cluster most features are licensed by the second position, only one, nasality, is anchored solely to the first — hence the name partial geminate for this type of cluster. (12) shows such a cluster.

(12) 

Following the algorithm given above, the complexity of a homorganic nasal + plosive cluster turns out to be 1 (the [nas] feature anchored to the first position).

The current proposal predicts that the representation of homorganic liquid + plosive and *s* + plosive clusters will involve two features independently anchored to the first consonantal position. Demonstrating whether this is indeed so, or whether the algorithm given for calculating the complexity of a consonant cluster is oversimplified is a task for future research.

If what has been said above is to stand, a model of consonantal representations must be sought according to which the glottal stop contains no features at all, the canonical places of articulation are composed of a single feature, while other places of articulation contain more than one feature. Without delving into the details of such models, let us note that it is not without reason that glottal consonants are considered placeless (Lass 1984, 179), in fact, the glottal stop is mere consonantalness (Szigetvári to appear). It is also noteworthy that in lenition processes affecting place of articulation the canonical plosives typically turn into a glottal stop — and not into each other —, while more complex places of articulation turn into the canonical types: e.g., Proto-Greek labiovelar plosives turn into labial and coronal plosives in Ancient Greek (cf. AG *penta* vs. Latin [k^{w}]*in*[k^{w}]*e* 'five'), while this place of articulation is retained in Latin, but reduced to velar in non-prevocalic position (*co*[k^{w}]*o* 'I cook' ~ *co*[k]*tus* 'cooked'). If lenition is seen as loss of features (e.g., Harris 1997), then plosives of canonical places of articulation can lose their only feature (becoming glottal), but cannot acquire another one (except through assimilation), hence a canonical plosive cannot turn into another canonical plosive without some external source of place of articulation. On the other hand, other places of articulation become canonical by losing some of their features (the one responsible for labiality or ve-

larity in the examples given above). Such phenomena provide evidence for the claim that the glottal stop is of complexity 0, canonical places of articulation are of complexity 1, while other places of articulation are more complex.

5. Conclusion

This paper offers a workaround for the situation bugging a number of researchers who notice that standard markedness metrics yield contradicting results. The solution proposed claims that markedness and complexity are almost the same, but not quite: the least complex categories in languages are often dispreferred and hence count as more marked than categories that are slightly more complex, but in other respects unmarked.

References

Blevins, Juliette 1995. The syllable in phonological theory. In: John A. Goldsmith (ed.): The handbook of phonological theory, 245–306. Blackwell, Cambridge MA & Oxford.

Charette, Monik 1992. Mongolian and Polish meet Government Licensing. In: SOAS Working Papers in Lingusitics and Phonetics 2 : 275–292.

Chomky, Noam – Morris Halle 1968. The sound pattern of English. Harper & Row, New York.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. In: John Kingston – Mary E. Beckman (eds): Papers in laboratory phonology I: Between the grammar and the physics of speech, 283–333. Cambridge University Press, Cambridge.

Cyran, Eugeniusz 2003. Complexity scales and licensing strength in phonology. Wydawnictwo KUL, Lublin.

Goldsmith, John A. 1990. Autosegmental and metrical phonology. Blackwell, Oxford & Cambridge MA.

Harris, John 1997. Licensing Inheritance: An integrated theory of neutralisation. In: Phonology 14 : 315–370.

Hume, Elizabeth 2003. Language specific markedness: The case of place of articulation. In: Studies in Phonetics, Phonology and Morphology 9 : 295–310.

Kaye, Jonathan – Jean Lowenstamm 1981. Syllable structure and markedness theory. In: Adriana Belletti – Luciana Brandi – Luigi Rizzi (eds): Theory of markedness in Generative Grammar, 287–315. Scuola Normale Superiore, Pisa.

Kaye, Jonathan – Jean Lowenstamm – Jean-Roger Vergnaud 1990. Constituent structure and government in phonology. In: Phonology 7 : 193–231.

Kean, Mary-Luise 1975. The theory of markedness in generative grammar. Doctoral dissertation, Massachusetts Institute of Technology.

Kirchner, Robert 2000. Geminate inalterability and lenition. In: Language 76:509–545.

Lass, Roger 1984. Phonology. An introduction to basic concepts. Cambridge University Press, Cambridge.

Lombardi, Linda 2002. Coronal epenthesis and markedness. In: Phonology 19:219–251.

Lombardi, Linda 2003. Markedness and the typology of epenthetic vowels. ROA 578-0203.

Maddieson, Ian 1984. Patterns of sounds. Cambridge University Press, Cambridge.

Prince, Alan – Paul Smolensky 1993. Optimality theory: Constraint interaction in generative grammar. Technical Report TR-2, Center for Cognitive Science, Rutgers University, New Brunswick, N.J. and Technical Report CU-CS-697-93, Department of Computer Science, University of Colorado, Boulder.

Prince, Alan S. 1984. Phonology with tiers. In: Mark Aronoff – Richard T. Oehrle (eds): Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students, 234–244. MIT Press, Cambridge MA & London.

Rebrus, Péter – Viktor Trón 2004. Fonetikai motiváció a fonológiai mintázatokban [The phonetic motivation of phonological patterns]. In: László Hunyadi (ed.): Kísérleti fonetika, laboratóriumi fonológia a gyakorlatban [Experimental phonetics and laboratory phonology: Practical aspects], 139–164. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen.

Rice, Keren 2003. Featural markedness in phonology: Variation. In: Lisa Cheng Rint Sybesma (eds): The second *Glot International* state-of-the-article book, The latest in linguistics, number 61 in Studies in Generative Grammar, 389–429. Mouton de Gruyter, Berlin & New York.

Rice, Keren to appear. Markedness in phonology. In: Paul de Lacy (ed.): The Cambridge handbook of phonology. Cambridge MA, Cambridge.

Siptár, Péter 2006. How to get rid of hiatuses. In: András Cser (ed.): Köszöntő kötet Szende Tamás tiszteletére [In honour of Tamás Szende], 59–70. Open Art, Budapest.

Szigetvári, Péter to appear. Manner as a skeletal relation. In: Friedrich Neubarth – Markus Pöchtrager – John Rennison (eds): Phonologica 2002: Structure and melody, proceedings of the 9th international phonology meeting, 1–3 November 2002 in Vienna. Mouton de Gruyter, Berlin & New York.

Trubetzkoy, N. S. 1969. Principles of phonology. University of California Press, Berkeley & Los Angeles.

Wilson, Colin 2001. Consonant cluster neutralisation and targeted constraints. In: Phonology 18:147–197.

Acta Linguistica Hungarica, Vol. 53 (4), pp. 449–466 (2006)
DOI: 10.1556/ALing.53.2006.4.4

ON THE ROLE OF VERBAL PARTICLES IN THE PROGRESSIVE IN HUNGARIAN*

KÁROLY VARASDI

Research Institute for Linguistics
Hungarian Academy of Sciences
Benczúr utca 33.
H-1068 Budapest
Hungary
varasdi@nytud.hu

Abstract: In this paper I discuss Hungarian progressive as it is expressed in focus-free sentences whose VP possibly contains a particle (verbal prefix). I define three simple distributional tests on the basis of which logical correspondences between certain types of expressions are established. These correspondences are then used to refute the hypothesis that the progressive in Hungarian is a stativizer. Finally, I take a broader look at the possibility of predicting the existence of the progressive reading in the case of particle plus verb complexes.

Keywords: progressive aspect, verbal particles, distributional tests, stativizer, aspectual composition

1. Introduction

This paper is about the Hungarian progressive as it is expressed in focus-free sentences whose VP possibly contains a particle (verbal prefix). In section 2 I define three distributional tests that will serve as the methodological backbone of the discussion. Each of these tests is a version of traditionally well-known tests but my concern will be to establish certain relations between the classes of Hungarian expressions compatible with

* I thank László Kálmán, Márta Maleczki, Anna Asbury and Zsófia Gyarmathy for their comments on this paper. Of course, none of them is responsible for any mistakes that may have remained in the text.

them. Based on the results, in section **3** I make a distinction between two types of progressive in Hungarian. This distinction is put to work in section **4** where I criticize Christopher Piñón's theory of the Hungarian progressive, and show that Hungarian provides strong counter-evidence to the view that the progressive is a stativizer. Finally, in section **5** I take a broader outlook at the aspectual semantics of verb particle complexes. The article ends with some suggestions for further work.

2. Three tests

In Hungarian, progressive expressions do not bear such simple morphosyntactic markers as their English counterparts. While in English the presence of *BE* plus *V-ing* signifies progressive interpretation, in Hungarian the nearest equivalent to that is a particular order of verb plus verbal particle accompanied by an even phonological stress pattern. However, as many Hungarian linguists have pointed out, this morphosyntactic pattern is only a sufficient (barring focus, see below) but not a necessary condition for a sentence to have a progressive interpretation.

A formally elaborated theory of the progressive in Hungarian can be found in Piñón (1995), and will be discussed in detail in section **4**. Based on Kiefer (1991; 1992a), Piñón introduces a battery of distributional tests which I will use here as well, though in a slightly generalized form. The terminology and tests in Piñón (1995) are as follows.

Definition 1 *A verbal expression E has an* ***event interpretation*** *iff it is compatible with time-span adverbials but not with durative adverbials:*[1]

(1) ⟨X⟩ ⟨E-zett⟩ öt perc alatt
⟨X⟩ ⟨E-d⟩ five minutes under
'⟨X⟩ ⟨E-d⟩ in five minutes'

[1] There are several ways to form durative adverbials in Hungarian beside the *öt percig* type, for example:

⟨X⟩ ⟨E-zett⟩ öt perc-en át/keresztül
⟨X⟩ ⟨E-d⟩ five minutes-on through/across
'⟨X⟩ ⟨E-d⟩ for five minutes'

As László Kálmán pointed out to me, *keresztül*, *át* and *-ig* differ in very subtle ways. However, I will ignore these minor differences in what follows and treat these phrases as completely equivalent forms expressing duration over a period.

(2) #⟨X⟩ ⟨E-zett⟩ öt perc-ig.
⟨X⟩ ⟨E-d⟩ five minutes-for.
'⟨X⟩ ⟨E-d⟩ for five minutes'

Definition 2 *An expression has a* ***process interpretation*** *iff it is compatible with durative adverbials, but not with time-span adverbials:*

(3) ⟨X⟩ ⟨E-zett⟩ öt perc-ig.
⟨X⟩ ⟨E-d⟩ five minutes-for.
'⟨X⟩ ⟨E-d⟩ for five minutes'

(4) #⟨X⟩ ⟨E-zett⟩ öt perc alatt
⟨X⟩ ⟨E-d⟩ five minutes under
'⟨X⟩ ⟨E-d⟩ in five minutes'

Definition 3 *An expression has a* ***progressive interpretation*** *iff it fits the following scheme:*

(5) ⟨X⟩ [éppen] ⟨E-zett⟩ [amikor...]
⟨X⟩ [just] ⟨E-d⟩ [when...]
'⟨X⟩ was ⟨E-ing⟩ when...'

Notice that the first two tests above are complex in the sense that they formulate conjunctive conditions. For logical reasons it is useful to disjoin these conjunctions and "factor out" their appropriate parts so that we can have logically "weaker" tests, out of which we can build up the tests used by Piñón, if required. These factors are as follows:[2]

Definition 4 *An expression has a* ***Type 1 interpretation*** *iff it is compatible with time-span adverbials:*

(6) ⟨X⟩ ⟨E-zett⟩ öt perc alatt
⟨X⟩ ⟨E-d⟩ five minutes under
'⟨X⟩ ⟨E-d⟩ in five minutes'

[2] In order to avoid unintended connotations, I introduce a neutral terminology.

Definition 5 *An expression has a* ***Type 2 interpretation*** *iff it is compatible with durative adverbials:*

(7) ⟨X⟩ ⟨E-zett⟩ öt perc-ig.
⟨X⟩ ⟨E-d⟩ five minutes-for.
'⟨X⟩ ⟨E-d⟩ for five minutes'

Finally, test **Type 3** tests the progressive interpretation:

Definition 6 *An expression has a* ***Type 3 interpretation*** *iff it fits the following scheme:*

(8) ⟨X⟩ [éppen] ⟨E-zett⟩ [amikor...]
⟨X⟩ [just] ⟨E-d⟩ [when...]
'⟨X⟩ was ⟨E-ing⟩ when...'

The above battery of tests can be used to characterize expressions in a straightforward manner. So, for example, we will say that expression E is of Type 1 iff E is admissible in the frame of Type 1, and we can represent this fact as $\mathsf{type}_1(E)$. Using this convention, we can characterize what is called a process interpretation by Piñón as follows: expression E has process interpretation iff $\mathsf{type}_1(E)$ is false but $\mathsf{type}_2(E)$ is true, and similarly for the event interpretation as well ($\mathsf{type}_1(E)$ is true but $\mathsf{type}_2(E)$ is false). Finally, the progressive interpretation can be represented simply as $\mathsf{type}_3(E)$.

In what follows, I will concentrate on VPs containing particles. Since the position of the particle in the sentence may depend on several factors that are orthogonal to the progressive aspect itself, I will confine the discussion to the simplest cases. In particular, I will only consider **focus-free sentences**, because the syntactic patterns of the progressive and the focus in Hungarian overlap, so it is expedient to treat them separately when one wants to concentrate on the progressive only.

I distinguish between two VP-types depending on the position of the verbal particle: in what I will call "prefixed VP" the particle comes immediately before the verb, and in what I will call "postfixed VP" it follows the verb. Following É. Kiss's (1998) model, we can assign the structures to them in Figures 1 and 2 respectively, though actually nothing will depend on this choice:

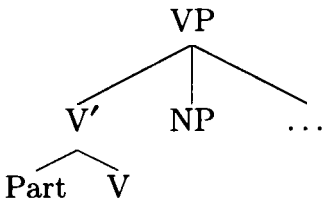


Fig. 1
Prefixed VP

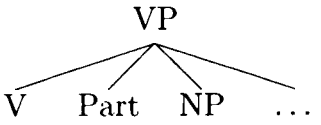


Fig. 2
Postfixed VP

3. Two types of progressive in Hungarian

On the basis of the foregoing discussion, some interesting observations can be made. First I state an important fact concerning the relationship between Type 2 and Type 3 expressions, which is as follows:

FACT 1

Type 2 $\subseteq$ Type 3

Expressions of Type 2 are also expressions of Type 3.

In other words, whenever an expression fits the test frame for Type 2, it also fits the frame for Type 3. This means that we predict that if, say, (9) is compatible with a durative adverbial like *fél órán át* ('for half an hour'), then it will pass the test for the progressive as well. And indeed, (10) is perfectly acceptable, as is (11).

(9) esett
rain-past-3sg

(10) Fél órán át esett
half hour through rain-past-3sg
'It rained for half an hour'

(11) Éppen esett, amikor...
just rain-past-3sg when...
'It was raining when...'

Here are some more examples to illustrate the generalization. The first sentence of the following pairs shows that a certain expression is admissible in the Type 2 frame, and the second that it is also admissible in the Type 3 frame:

(12) János és Mari másfél órán át sakkoztak
John and Mary one-and-a-half hour through play-chess-past-3pl
'John and Mary played chess for one hour and a half'

(13) Mikor telefonáltam, János és Mari éppen sakkoztak
when phone-past-1sg John and Mary just play-chess-past-3pl
'When I called them, John and Mary were playing chess'

(14) János negyed órán keresztül krumplit hámozott
John quarter-of-an-hour through potato-acc peel-past-3sg
'John peeled potatoes for a quarter of an hour'

(15) Amikor megérkeztem, János éppen krumplit hámozott
when arrive-past-1sg John just potato-acc peel-past-3sg
'When I arrived, John was peeling potatoes'

(16) Robi néhány percig a másik part felé úszott
Rob some minute-for the other bank toward swim-past-3sg
'Rob swam toward the other bank for a couple of minutes'

(17) Amikor megpillantottam, Robi éppen a másik part felé úszott
when see-him-past-1sg Rob just the other bank toward swim-past-3sg
'When I saw him, Rob was swimming toward the other bank'

Fact 1 is a distributional generalization. The corresponding **semantic** generalization could be stated as a constraint on atelic expressions in Hungarian: they are admissible in both Type 2 and Type 3. Indeed, this hypothesis is borne out further by the observation that stage-level state expressions in Hungarian also belong to Type 2 **and** Type 3:

(18) Mari két hétig beteg volt
Mary two weeks-for ill be-past-3sg
'Mary was ill for two weeks'

(19) Amikor felhívtam, Mari éppen beteg volt
when call-her-past-1sg Mary just ill be-past-3sg
'When I called her, Mary was ill'

However, there are Type 3 VPs that are **not** Type 2.[3] The following examples contain VPs which are of Type 3 but not Type 2 (note that all of the examples below contain postfixed VPs):

(20)#Éva pár percig ment le a pincébe
Eve a-couple minute-for go-past-3sg prt the cellar-into

(21) Mikor összefutottunk, Éva éppen ment le a pincébe
when meet-past-1pl, Eve just go-past-3sg prt the cellar-into
'When we met, Eve was going down into the cellar'

(22)#Kata néhány másodpercig vette fel a kabátját
Kate some second-for put-on-past-3sg prt the coat-poss-acc

(23) Amikor észrevettem, Kata éppen vette fel a kabátját
when notice-her-past-1sg Kate just put-on-past-3sg prt the coat-poss-acc
'When I noticed her, Kate was putting on her coat'

In other words, Type 3 is not equal to Type 2. We can represent the relevant relationship between the appropriate sets of expressions as proper inclusion:

FACT 2

Type 2 $\subsetneq$ Type 3

The set of expressions of Type 2 is properly included in the set of expressions of Type 3.

Let us call those expressions that belong to Type 3, but not to Type 2, **proper Type 3** expressions:

$$\textbf{Proper Type 3} = \textbf{Type 3} \setminus \textbf{Type 2}$$

Now we can state an important generalization (which we hinted at above) concerning postfixed VPs:

FACT 3

Postfixed VPs are proper Type 3

All postfixed VPs are proper Type 3 VPs.

[3] This important fact, which we will use in section **4.2**, was first noted by Kiefer (1991).

At this point a caveat is in order: the appearance of the verbal particle behind the verb can also be the result of the presence of focus as in (24).

(24) JÁNOS írta meg a levelet
John-foc write-past-3sg prt the letter-acc
'It was John who wrote the letter'

But this particular VP is not possible as a progressive VP:

(25) #János éppen írta meg a levelet, amikor...
John just write-past-3sg prt the letter-acc when...

I will return to the question of predicting whether a VP is a possible progressive VP or not in section **5**.

Note that I am not claiming that postfixed VPs exhaust the class of proper Type 3 expressions. Indeed, there are well-formed expressions belonging to Type 3 that contain no particle at all, as in (26) which is a well-formed progressive sentence, although (27) is not acceptable. In what follows, however, I will ignore these instances of Type 3, and only concentrate on VPs containing particles.

(26) Éva éppen a folyóhoz biciklizett, amikor...
Eve just the river-to cycle-past-3sg when...
'Eve was cycling to the river when...'

(27) #Éva néhány percig a folyóhoz biciklizett
Eve some minute-for the river-to cycle-past-3sg

4. Piñón's theory of the progressive in Hungarian

In this section I discuss an important theory of the Hungarian progressive that is a worked-out formal proposal both from a syntactic as well as a semantic point of view. This is not to say, of course, that scholars had not dealt with the question earlier. On the contrary, Kiefer (1982; 1991; 1992a) as well as É. Kiss (1987; 1992) had looked at the topic from both a descriptive and a theoretical angle. However, it was Christopher Piñón in Piñón (1995) who proposed a unified formal theory of the progressive in Hungarian that pays equal attention both to its syntax and its semantics.

The first part of Piñón's article is polemic and aims at refuting the particular syntactic theory of the progressive operator presented in

É. Kiss (1987; 1992). He also takes a critical look at Kiefer's earlier analyses, though he seems to agree with Kiefer on several important points. In what follows I will examine the **semantic** side of Piñón's theory in details.

The principal claim of the article is expressed in Piñón's words as

> "The heart of my analysis is to postulate a difference in how progressives of process and event expressions are formed. Progressives of process expressions can be formed directly; progressives of event expressions, in contrast, cannot. Event expressions must first be converted into process expressions before a progressive can be formed. The conversion of event expressions into process expressions is overtly marked in Hungarian: it is what requires verb movement." (Piñón *op.cit.*, 162)

This approach to the progressive is motivated by Moens-Steedman (1988) and the idea of the progressive as a "coercive operator" defended therein, cf. Piñón (*op.cit.*, 169). Piñón lists three assumptions that guide his analysis. The first two postulates are of lesser importance to us now, but the following one is crucial:

> "PROG [a semantic operator—K. V.] takes only process predicates as input. In order for PROG to apply to event predicates, the latter first have to be converted into process predicates. The semantic operator PR(ocess) converts event predicates into process predicates. The morphosyntactic representation of PR is [Pr]. [Pr] also does not fill the preverbal focus position." (*ibid.*, 168)

These claims can be represented schematically as constraints on the typing of the PROG and PR operators as follows:

PROG(*Process*);
PROG(PR(*Event*)), where
PR: *Event* → *Process*

Piñón argues in the paper that it is a mistake to locate the PROG operator (or rather, its morphosyntactic realization) in the [Spec, VP] position, as claimed by É. Kiss. Instead, he suggests a different position for it (as an adjunct of a particular X′-projection). I am not going to discuss the details of this suggestion because it basically pertains to the particular syntactic framework the analysis is couched in. Rather, I want to concentrate first on how plausible it is to suppose the existence of PR in Hungarian in the way Piñón suggests, and then I will discuss the deeper semantic assumptions behind the analysis Piñón offers.

4.1. The existence of PR

As we have seen, Piñón claims that PROG can only take process predicates as input. He also claims that "the PR operator, which 'coerces' event predicates into process predicates, effectively imposes the syntactic order of verb plus PV [particle—K. V.]" (Piñón *op.cit.*, 169), and also that "the semantic operator PR applies only if we need to convert an event expression into a process expression [...]. If we have a process expression to begin with, then PROG can apply without the mediation of PR." (*ibid.*, 178).

First note that this suggestion has the merit of explaining why it is always possible to use process expressions in the progressive. As the input of PROG must be a process, feeding a process predicate into it is always possible, as is in fact borne out by the data. But PR is only invoked when PROG needs it to convert an event predicate into a process predicate; otherwise it does not do anything. In other words, in the case of event expressions PR operates if and only if PROG does (and then it imposes the verb plus particle order), and when the input is a process expression, it does nothing. But this prompts the question immediately: What other theory-independent reasons are there for postulating such an operator?

The assumption that calls for postulating PR is that the progressive requires a process as input. This hypothesis is already present in Moens and Steedman's transition diagram (Moens–Steedman 1988, 18), where the authors claim that the progressive operator works on processes and maps them to states (the so-called progressive states).[4] However, there is no convincing evidence supporting the hypothesis that the input of the progressive operator must be a process. On the contrary, this assumption is questionable even in English. There is a class of verbs in English as well as in Hungarian that are stative and still can appear in the progressive as in (28).

(28) János éppen aludt /állt, amikor...
John just sleep-past-3sg /stand-past-3sg, when...
'John was sleeping/standing when...'

These "dynamic state" verbs are undoubtedly static (they do not imply any change), still they are perfectly admissible in the progressive in both

[4] This idea has its predecessor in Vlach (1981) where it is claimed that the progressive is a sort of "stativizer"; see below.

languages. But without this assumption the support for postulating PR seems to disappear.

Later in the analysis Piñón draws upon another assumption that I already mentioned above, namely, that the progressive is a stativizer:

> "The output of PROG is a state predicate; this is consistent with the standard view that progressives describe states (e.g., see Asher (1992) for a recent example)." (Piñón *op.cit.*, 180)

But, at least in Hungarian, there is strong evidence that this assumption is also false.

4.2. The progressive is *not* a stativizer in Hungarian

The claim that the progressive is a stativizer was first suggested by Vlach (1981), and has been accepted by many linguists since. The claim is as follows:

(29) **The progressive as a stativizer**
Whenever the progressive is true there exists a state, the progressive state, which holds as long as the progressive is true.

Let us turn to statives now. Stative expressions, similarly to Type 2 VPs, are usually compatible with durative temporal modifiers in Hungarian, as we have already seen (e.g., sentence (18)). Some further examples are:

(30) János Londonban élt húsz évig
John London-in live-past-3sg twenty years-for
'John lived in London for twenty years'

(31) A könyv fél órán át az asztalon hevert
the book half-an-hour through the table-on lie-past-3sg
'The book lay on the table for half an hour'

This compatibility only disappears when it is hard or actually impossible to find a particular event responsible for the transition into and/or out of the state described by the stative expression, either because the boundaries of the state in question are fuzzy (as in (32)), or because it is constitutive of the subject (as in (33)), or because the state is irreversible (as in (34)):

(32) #János két hétig tudta a választ.
John two weeks-for know-past-3sg the answer
'John knew the answer for two weeks'

(33) #Péter hatvan évig ember volt
Peter sixty years-for human be-past-3sg
'Peter was a human being for sixty years'

(34) #Mari száz évig halott volt
Mary one-hundred years-for dead be-past-3sg
'Mary was dead for one hundred years'

The following fact, which we will use later on, is generally true of states:

FACT 4

From points to intervals

If a state is predicable of every point in an interval, then it holds throughout the interval.

Also, stative expressions can be modified by durative adverbials measuring the temporal duration of the state:

FACT 5

Modification of states by durative adverbials

If a state holds throughout an interval, the VP describing it can be modified by a durative adverbial specifying the length of the interval.

For example, if (35) was true exactly in each and every moment of the interval [17:00,17:05], then (36) is also true.

(35) A könyv az asztalon van
the book the table-on be-pres-3sg
'The book is on the table'

(36) A könyv öt percig az asztalon volt
the book five minute-for the table-on be-past-3sg
'The book was on the table for five minutes'

Now let us take the following sentence containing a postfixed VP (the particle *le* literally means 'down'):

(37) Jóska éppen vitte le a bort a pincébe, amikor...
Joe just take-past-3sg prt the wine-acc the cellar-to, when...
'Joe was carrying the wine down to the cellar when...'

Clearly, Joe's carrying the wine down to the cellar is a well-defined event: it has a definite beginning (he takes the first step toward the door of the cellar) and, if succeeds, has a definite end (he steps in the cellar and puts down the wine). But even if he does succeed in completing the event, the following sentence is odd:

(38) #Jóska fél percig vitte le a bort a pincébe
Joe half minute-for take-past-3sg prt the wine-acc the cellar-to
Intended meaning: 'Joe's carrying the wine down to the cellar took half a minute'

This is not suprising: we have already seen on page 455 that proper Type 3 VPs, which are first-class progressives, are not compatible with durative adverbials. But if the progressive is a stativizer, then this is quite unexpected.

Note that this problem cannot be solved by claiming that the progressive can only be true at moments and, as durative adverbials require a non-null interval, the incompatibility is therefore explained. For suppose that a progressive sentence is actually true at each and every moment of a particular interval I. Then, if the progressive is a stativizer, the progressive **state** is also true at each and every moment of I. But when a state can be predicated of each and every moment of an interval then, by Fact 4, it holds throughout the interval. Then, by Fact 5, the VP is modifiable by a durative adverbial. But this is not so, as sentence (38), or any other similar sentence having a proper Type 3 VP, illustrates.

Actually, the existence of proper Type 3 progressive expressions provides a strong counter-argument to the claim that the progressive operator in Hungarian is a stativizer. For again suppose that it is, and that it pertains to **non-null** intervals. Now let a proper progressive sentence be predicated of a particular non-null interval J. Then, if the progressive is a stativizer, a progressive **state** can also be predicated of J. But when a state can be predicated of an interval, then, by Fact 5, its description is compatible with a durative adverbial measuring the length of the interval in question. But as we have seen, proper progressive expressions are **never** compatible with durative adverbials.

Putting these facts together, we can conclude that we have a contradiction both when we assume that proper Type 3 (i.e., proper progressive)

expressions can be predicated of null-intervals (moments) and when we assume that they can be predicated of non-null intervals. Since we have exhausted the logical possibilities, we must conclude that the existence of proper Type 3 progressives provides counter-evidence to the claim that the progressive is a stativizer in Hungarian.

5. On the aspectual properties of particle–verb complexes

In this section I take a broader look at the aspectual properties of particle–verb complexes. Let me start with a generalization concerning the relationship between proper Type 3 and Type 1 expressions: Whenever a proper Type 3 VP is acceptable (and then it is in the progressive aspect), then there exists a corresponding VP of Type 1 in which the same particle is in the preverbal position. For example, look at the following sentence:

(39) Réka ment be a könyvtárba
Réka go-past-3sg prt the library-into
'Réka was going into the library'

Putting the particle *be-* 'into' in front of the verb *ment* 'went' results in (40) which is an instance of Type 1.

(40) Réka be-ment a könyvtárba
Réka into-go-past-3sg the library-into
'Réka went into the library'

More precisely,

FACT 6

From proper Type 3 to Type 1

If a proper Type 3 VP is well-formed, then there is a corresponding Type 1 VP, but **not** vice versa.

Although in some of these cases we might talk about the "movement" of the particle to the front of the verb, there is nothing to move in those cases when the order of verb plus particle is not possible and still the order of particle plus verb is. For example, as we have seen earlier, in (41), the particle *meg* cannot appear after the verb in a neutral sentence to form a progressive VP:

(41) Írta a levelet
write-past-3sg the letter-acc
'(S)he was writing the letter'

(42) #Írta meg a levelet
write-past-3sg prt the letter-acc

However, putting *meg* in front of the verb is possible and makes a Type 1 VP:

(43) Meg-írta a levelet
prt-write-past-3sg the letter-acc
'(S)he has written/wrote the letter'

What explains Fact 6? A semantic explanation might go like this. The existence of the progressive interpretation signifies that the expression is an (atelic) process/dynamic state or a (telic) accomplishment. Atelic VPs belong to Type 2 (and, as we have seen, also to Type 3), while proper Type 3 contains accomplishments. Since postfixed VPs are all proper Type 3 expressions, a postfixed VP is an accomplishment in the progressive. Since all progressive accomplishments can, in principle, be accomplished, we predict the existence of a form that expresses the completion of the accomplishment in question.[5]

A simple compositional picture emerges on the basis of these facts if we assume that in Hungarian the progressive is more basic than the perfective (the latter being marked by filling out the preverbal position with the particle).[6] The meaning contribution and the aspectual contribution of the particle must be separated because they are orthogonal. The meaning contribution of the particle does not depend on the surface configuration of the sentence (i.e., whether progressive or perfect). For example, whatever the meaning that the particle *le* contributes to the following sentences is, it is the same in both cases:

(44) János éppen olvasta le a vízórát, amikor...
John just read-past-3sg prt the water-meter, when...
'John was reading off the water-meter when...'

[5] This line of thought, being semantic in nature, can be extended to all expressions of proper Type 3, like the one in sentence (27); however, as mentioned above, in this paper I am concentrating on VPs containing a particle.

[6] I am ignoring any other material that may appear in this position and concentrate on verbal particles exclusively.

(45) János le-olvasta a vízórát.
John prt-read-past-3sg the water-meter.
'John read off/has read off the water-meter.'

The difference is only aspectual.

The above example would be analyzed by most linguists as an instance of a non-compositional compound: *le* and *olvas* together make a meaning that is unpredictable on the basis of their "literal meanings". The point of this example is that the meaning of the verb plus particle complex is often unpredictable, so it should be taken semantically as one unit. However, the **position** of the particle has a well-defined aspectual contribution which can be summed up as follows: if it is in front of the verb, it makes the sentence perfective, while when it follows the verb, the sentence will be (proper) progressive.

These considerations prompt the question whether it is possible to predict on the basis of the meaning of the particle and the verb if the order of verb plus particle (i.e., the progressive interpretation) is possible. In some cases this is certainly possible. For example, when the verb expresses continuous spatial motion (e.g., *rohan* 'run') and the particle expresses a spatial direction (e.g., *ki* 'out'), the complex expression can be made progressive:

(46) János éppen rohant ki az állomásra, amikor...
John just run-past-3sg prt the station-to when...
'John was running to the railway station when...'

However, it seems that, apart from these cases, the chances of such a prediction are small, and the reason is hard to identify in full generality. For example, Kiefer (1991) makes the (tentative) hypothesis that the impossibility of certain morphosyntactically expressed progressives might be due to the fact that their particle**less** version already expresses the required progressive meaning (*op.cit.*, 265). This blocking hypothesis could explain why (47) is not well formed: simply because the sentence in (48) exists and expresses the required meaning.

(47) #János éppen olvasta el a *Háború és békét*
John just read-past-3sg prt the war and peace-acc
Intended meaning: 'John was reading *War and Peace*'

(48) János éppen olvasta a *Háború és békét*
John just read-past-3sg the war and peace-acc
'John was reading *War and Peace*'

However, this blocking strategy breaks down when it comes to explaining why (49) is fine.

(49) János éppen kente be a padlót viasszal
John just smear-past-3sg prt the floor-acc wax-with
'John was smearing the floor with wax'

Indeed, if we leave out the particle from the above sentence, the resulting sentence is perfectly acceptable and has a progressive meaning:

(50) János éppen kente a padlót viasszal
John just smear-past-3sg the floor-acc wax-with
'John was smearing the floor with wax'

One might object that the two progressive sentences are not completely synonymous, that is, (49) and (50) have slightly different meanings: (49) strongly suggests that John had the intention to cover the whole of the floor with wax, and this element is missing from (50). However, even if this is so, we cannot explain what forbids the insertion of *el* in (47) to express a similar intention on John's part to read the whole of the novel. Note also that claiming that *el* somehow, by force of its meaning, makes *olvasta a Háború és békét* perfective and this is why it cannot occur in the progressive sentence does not solve the problem either, because the meaning contribution of *be* in *be-kente a padlót viasszal* is similar or even identical to that of *el* in *el-olvasta a Háború és békét*. In fact, the contribution of the particles in both cases can be paraphrased as "intentionally continues to do the action until the whole of the object has been subjected to it".

6. Conclusion and further work

We have seen that the precise description of the language-dependent set-theoretical relations between sets of expressions compatible with well-chosen distributional tests can help in establishing such theoretical results as the progressive is not a stativizer in Hungarian. Therefore, it would be important to extend the scope of such empirical investigations to cover

a larger variety of expressions in Hungarian. On the purely theoretical side, investigations into the semantics of temporal adverbials used in such tests are also needed. Of course, these tasks are strongly interdependent, and should be carried out simultaneously.

References

Asher, Nicholas 1992. A default, truth conditional semantics for the progressive. In: Linguistics and Philosophy 15 : 469–508.

É. Kiss, Katalin 1987. Configurationality in Hungarian. Akadémiai Kiadó, Budapest.

É. Kiss, Katalin 1992. Az egyszerű mondat szerkezete [The structure of clauses]. In: Kiefer (1992b, 79–176).

É. Kiss, Katalin 1998. Mondattan [Syntax]. In: Katalin É. Kiss – Ferenc Kiefer – Péter Siptár: Új magyar nyelvtan [A new Hungarian grammar], 15–184. Osiris Kiadó, Budapest.

Kiefer, Ferenc 1982. The aspectual system of Hungarian. In: Ferenc Kiefer (ed.): Hungarian and general linguistics, 293–329. John Benjamins, Amsterdam & Philadelphia.

Kiefer, Ferenc 1991. A progresszív aspektus a magyarban [The progressive aspect in Hungarian]. In: Magyar Nyelv 87 : 257–77.

Kiefer, Ferenc 1992a. Az aspektus és a mondat szerkezete [Aspect and syntactic structure]. In: Kiefer (1992b, 797–886).

Kiefer, Ferenc (ed.) 1992b. Strukturális magyar nyelvtan 1. Mondattan [A structural grammar of Hungarian, Vol. 1: Syntax]. Akadémiai Kiadó, Budapest.

Moens, Marc – Mark Steedman 1988. Temporal ontology and temporal reference. In: Computational Linguistics 14 : 15–29.

Piñón, Christopher 1995. Around the progressive in Hungarian. In: István Kenesei (ed.): Approaches to Hungarian 5 (Levels and structures), 153–90. JATEPress, Szeged.

Vlach, Frank 1981. The semantics of the progressive. In: Philip Tedeschi – Annie Zaenen (eds): Tense and aspect. Syntax and semantics 14, 271–292. Academic Press, New York.

Acta Linguistica Hungarica, Vol. 53 (4), pp. 467–481 (2006)
DOI: 10.1556/ALing.53.2006.4.5

DIRECT AND INDIRECT SPEECH IN STRAIGHT-TALKING ISRAELI

GHIL'AD ZUCKERMANN

The University of Queensland
St Lucia QLD 4072
Australia
gz208@cam.ac.uk

Abstract: Israeli is currently one of the official languages of the State of Israel. It is a fusional synthetic language, with non-concatenative discontinuous morphemes realised by vowel infixation. This typological paper demonstrates that there is a clear distinction in Israeli between direct and indirect speech. The indirect speech report, which is a subset of complement clauses, is characterized by a shift in person, spatial and temporal deixis. However, unlike in English, the verbs usually do not undergo a tense shift. Israeli has various lexicalized direct speech reports. By and large, Israeli reported speech constructions reflect Yiddish and Standard Average European patterns, often enhancing a suitable pre-existent Hebrew construction.

Keywords: Hebrew, reported speech, Yiddish, lexical derivation, Congruence Principle

1. Introduction

1.1. General information

Israeli (Zuckermann 1999; 2006b, a.k.a. "Modern Hebrew") is currently one of the official languages — with Arabic and English — of the State of Israel, established in 1948 on 20,770 km^2 (0.22 of Hungary) in the Middle East. It is spoken to varying degrees of fluency by its 7,026,000 citizens (as of May 2006) — as a mother tongue by most Jews (whose total number exceeds 5.6 million), and as a second language by Muslims

1216–8076/$ 20.00

(Arabic-speakers), Christians (e.g., Russian- and Arabic-speakers), Druze (Arabic-speakers) and others.

1.2. Grammatical profile

Israeli is a fusional synthetic language, with non-concatenative discontinuous morphemes realised by vowel infixation. Compare (1) and (2), both formed from the root *p.ṭ.r.*, but fitted into two distinct verb-templates:

(1) נפטר
niftár
pass.away:3msg.past
'(he) passed away'

(2) התפוטרנו
hitputárnu
resign:1pl.past:"coercive"
'We "resigned" (implying that we were encouraged to do so; had we not done so, we would have been fired anyway).'

Israeli is a head-marking language. It is nominative-accusative at the syntactic level and partially also at the morphological level. As opposed to Biblical Hebrew — whose constituent order is VAO(E)/VS(E) — but like Standard Average European and English, the usual constituent order of Israeli is AVO(E)/SV(E). Thus, where there is no case marking, one can resort to the constituent order.

The main clause in Israeli consists of (a) clause-initial peripheral markers, e.g., discourse markers; (b) NP(s) or complement clause(s); (c) a predicate — either verbal, copular or verbless; (d) clause-final peripheral elements, e.g., discourse markers. The only obligatory element is the predicate, e.g., *higáti* 'arrive:1sg.past'.

Sentences (3), (4) and (5) are examples of a verbal, copular and verbless clause, respectively:

(3) הילדה אכלה תפוח.
[ha-yaldá]$_A$ [akhl-á]$_V$ [tapúakh]$_O$
[def-girl]$_A$ [eat:3past-fsg]$_V$ [apple]$_O$
'The girl ate an apple.'

(4) הילדה הזאת היא אחות שלי.
[ha-yaldá ha-zòt]$_{CS}$ [hi]$_{COP}$ [akhót shel-ì]$_{CC}$
[def-girl def-prox.fsg]$_{CS}$ [cop:fsg]$_{COP}$ [sister gen-1sg]$_{CC}$
'This girl is my sister.'

(5) הילדה חכמה.
[ha-yaldá]$_{VCS}$ [khakham-á]$_{VCC}$
[def-girl]$_{VCS}$ [clever-f]$_{VCC}$
'The girl is clever.'

There are various types of subordinate clause, e.g., adverbial (denoting comparison, time, place, condition, concession, reason, result, goal, state), adjectival/relative, and nominal/ complement. By and large, these follow the Standard Average European profile. Indirect speech report fits into the overall system of complement clauses—see section **3.1.**

2. Speech report constructions

2.1. The DSR/ISR distinction

Due to (*inter alia*) the lack of evidentials in the language, Israeli does not possess any monoclausal speech report construction. It has a clear distinction between multiclausal direct speech report (henceforth, DSR) and multiclausal indirect speech report (henceforth, ISR), the ISR generally being more common than DSR.

Distinguishing features characterizing ISR:

(i) Shift in person deixis, e.g., 2 > 1
(ii) Shift in spatial and temporal deixis, e.g., 'today' > 'that day', also spatial demonstratives, e.g., 'this' > 'that'
(iii) Obligatory presence of a complementizer immediately before the speech report, unless the speech report is of the 'infinitive' (iv) or interrogative type (v)
(iv) In report of commands: imperative/future verb > "infinitive" (tenseless verb, commonly referred to in Israeli grammar as "infinitive", thus, henceforth, INF)
(v) In report of questions: interrogative-less yes/no question > *im* (lit. 'if') or *ha-ím* (lit. 'INTER-if') 'whether' **immediately before** the indirect question

Unlike in English, verbs do not undergo a tense shift. The future tense in a speech act in the past does not become future-in-the-past in ISR (see (16)); unlike English (cf. *would*), Israeli does not have a form of future-in-the-past. Similarly, present tense in a speech act in the past usually does not become past in ISR (see (9)).

Like in English, in indirect "*wh*-questions", the interrogative remains the same. Unlike in English, there is no change in constituent order (see (15)–(16)).

Distinguishing features characterizing DSR:

(i) Special intonation contour, tending to be mimetic
(ii) Possible intonational break before the DSR
(iii) Possible absence of a reporting verb
(iv) Possible discontinuity of the DSR
(v) Possible use of a vocative particle
(vi) In writing: presence of quotation marks, as well as exclamation/question mark

Sentences (6) and (7) demonstrate the shift in person deixis:

(6) הם אמרו לנו: "אין לנו כסף!"
hem amr-ú l-anù [én l-**anù** késef]$_{DSR}$
3mpl say:3past-pl dat-1pl [exis.cop:neg dat-**1pl** money]
'They told us: "We have no money!"'

(7) הם אמרו לנו שאין להם כסף.
hem amr-ú l-anù ⟨**she**-én l-**ahèm** késef⟩$_{ISR}$
3mpl say:3past-pl dat-1pl ⟨**comp**-exis.cop:neg dat-**3mpl** money⟩
'They told us that they had no money.'

Sentences (8) and (9) demonstrate the lack of tense shift in verbs:

(8) טלי לחשה: "אני רוצה גלידה".
Tali lakhash-á [anì **rotsá** glída]$_{DSR}$
Tali whisper:3past-fsg [**1sg want:fsg.pres** ice.cream]
'Tali whispered: "I want ice cream!"'

(9) טלי לחשה שהיא רוצה גלידה.
Tali lakhash-á ⟨**she**-hì **rotsá** glída⟩$_{ISR}$
Tali whisper:3past-fsg ⟨**comp-3fsg want:fsg.pres** ice.cream⟩
'Tali whispered that she wanted ice cream.'

Note that—in contrast to the English translation –the verb 'want' is in the present tense in both sentences.

2.2. Exclamative and vocative DSR

Obviously, not **every** DSR is transformable into an ISR:

(10) כש"גילו" למדלן אולברייט שהיא יהודייה, היא צווחה: "אוי ויי!"
k-she-"gil-ú" le-mádlen ólbrayt ⟨she-hí yehudi-yá⟩$_{ISR}$
when-comp-"reveal":past-3pl to-Madeleine Albright ⟨comp-3fsg Jewish-fsg⟩
hi tsavkh-á [óy véy!]$_{DSR}$
3fsg scream:3past-fsg [Oy vey]
'When it was "revealed" to Madeleine Albright that she was Jewish, she screamed: "Oy vey!"'

ISR cannot convey the associations accompanying a DSR vocative particle:

(11) הוא צעק עליה: "תפסיקי לנדנד, יא נודניקית!"
hu tsaák al-èa:
3msg shout:3msg.past on-3fsg
[tafsík-i le-nadnéd, **ya** núdnik-it!]$_{DSR}$
[stop:2fut/imp-fsg inf-bother **voc:derog** pest-fsg]
'He shouted at her: "Stop bothering, *ya* pest!"'

The vocative particle *ya* (cf. archaic English *O*, as well as contemporary colloquial (Antipodean) English *ya*, or *y'*, from *you*) is currently derogatory in the sense that it only precedes derogatory NPs. This particle can be traced back to the vocative exclamatory Arabic particle يا [ya:]. Initially, Israeli *ya*—just like in Arabic—was not derogatory—see the Israeli songs *ya mishlatí* 'O my fortified cliff' and *ya khabíbi* 'O my dear'. However, native Israeli-speakers are aware of the Arabic etymon and—perhaps due to the negative (e.g., terroristic) associations of Arabic among Israelis—*ya* underwent semantic—or rather pragmatic—narrowing: pejoration.

2.3. ISR complementizer

As shown in (7) and (9), ISR usually uses the common Israeli complementizer *she* [ʃe] 'that', which—just like English *that*—also acts as a

relativizer. *She-* 'that' can be traced back to the Hebrew complementizer *she-* 'that', which derives from the Hebrew relativizer *she-* 'that'. One etymological analysis is that *she-* is a shortened form of the Hebrew relativizer *'asher* 'that', which is related to Akkadian *'ashru* 'place' (cf. Semitic **'athar*).

Instead of using the *she-* complementizer, a more formal Israeli writer could use the rare complementizer *ki* 'that', which derives from the Hebrew complementizer *kī* 'that', from *ki* 'because'. Consider the following minimal pair:

(12) הנאשם הצהיר כי הוא חף מפשע.
ha-neeshám hitsír ⟨**ki** hu khaf mi-pésha⟩$_{\text{ISR}}$
def-accused:msg declare:3msg.past ⟨**comp** 3msg clean from-crime⟩
'The accused declared that he was innocent.'

(13) הנאשם זוכה כי הוא חף מפשע.
ha-neeshám zuká [**ki** hu khaf mi-pésha]$_{\text{CAUS}}$
def-accused:msg acquit:3msg.past:pass [**caus** 3msg clean from-crime]
'The accused was acquitted because he was innocent.'

Whereas in (12) *ki* introduces an ISR, in (13) it introduces a causal clause. But such versatility can easily result in ambiguity:

(14) הן לא סיפרו לי כי כבר הסבירו את זה.
hen lo sipr-ú l-i
3fpl neg tell:3past-pl dat-1sg
⟨**ki** kvar hisbír-u et ze⟩$_{\text{ISR/CAUS}}$
⟨comp/caus already explain:3past-pl acc prox.msg⟩
'They (f) did not tell me **that** it had already been explained.'
or 'They (f) did not tell me (about it) **because** it had already been explained.'

Thus, *ki* is often avoided even by Israelis attempting to write in a high-flown manner. As opposed to *she-*, I categorize *ki* as a prescriptive complementizer *tout court.* That said, some French-speaking immigrants to Israel use the complemetizer *ki* less rarely than other Israelis because of the serendipitous phonetic similarity to the French complementizer *que* 'that'—cf. Zuckermann (2006b).

2.4. Direct and indirect questions

Unlike in English, Israeli indirect questions demonstrate neither a shift in verb tense nor a change in constituent order. Thus, besides the distinct intonation, the only difference between (15) and (16) is the pronominal suffix used with the genitive *shel* 'of' (i.e., the only shift is in person):

(15) הוא שאל אותה: "מתי ההורים שלך יגיעו?"
hu shaál ot-à
3msg ask:3msg.past dat-3fsg
[**matáy** ha-hor-ím shel-àkh **yagí-u**]$_{DSR}$
[**when** def-parent-mpl gen-**2**fsg **arrive:3fut-pl**]
'He asked her: "When will your parents arrive?"'

(16) הוא שאל אותה מתי ההורים שלה יגיעו.
hu shaál ot-à
3msg ask:3msg.past dat-3fsg
⟨**matáy** ha-hor-ím shel-à **yagí-u**⟩$_{ISR}$
⟨**when** def-parent-mpl gen-**3**fsg **arrive:3fut-pl**⟩
'He asked her when her parents would arrive.'

Thus, one may regard the indirect question as a semi-direct speech report.

Although Standard Average European (often via Yiddish) is undoubtedly an important source for Israeli reported speech, this indirect question construction seems to have already existed in Hebrew too. Consider, for example, Biblical Hebrew *lo noda'* ⟨*mi hikkáhu*⟩ 'it be not known ⟨who hath slain him⟩' (Deuteronomy 21 : 1). Such multiple causation corresponds with the Congruence Principle: if a feature exists in more than one contributor, it is more likely to persist in the target language (see Zuckermann 2003).

2.5. Direct and indirect commands: infinitive ISR

ISR can lack a complementizer and instead begin with a tenseless verb, commonly referred to as "infinitive".

(17) המפקדת הורתה: "תהיו כאן תוך שבע דקות!"
ha-mefakéd-et hortá [**tiyú** kan tokh shéva dak-ót]$_{DSR}$
def-commander-fsg order:3fsg.past [**be:2pl-fut/imp** here within seven minute-pl]
'The commander (f) ordered: "Be here within seven minutes!"'

(18) המפקדת הורתה להיות שם תוך שבע דקות.
ha-mefakéd-et hortá ⟨**li-yót** sham tokh shéva dak-ót⟩$_{ISR}$
def-commander-fsg order:3fsg.past ⟨**inf-be** there within seven minute-pl⟩
'The commander (f) ordered to be back there within seven minutes.'

DSR has the capacity to be more nuanced than ISR. Thus, one ISR can be parallel to several distinct DSRs. For example, (18) can be the ISR not only of (17) but also of (19), which includes a semantic future perfect, realised morphologically in colloquial Israeli by the past:

(19) המפקדת הורתה: "תוך שבע דקות הייתם כאן!"
ha-mefakéd-et hortá [tokh shéva dak-ót **haítem** kan]$_{DSR}$
def-commander-fsg order:3fsg.past [within seven minute-pl **be:2mpl.past** here]
'The commander (f) ordered: "Within seven minutes, you will have arrived back here!"'

3. Syntactic role of speech report content

3.1. ISR versus complement clause

Israeli ISR conforms to complement clause structure. The following three sentences, which constitute a continuum, demonstrate that ISR is a subset of complementation:

(20) אני יודע שהיא יפה.
anì yodéa ⟨she-hì yaf-á⟩$_{COMP}$
1sg know:msg.pres ⟨comp-3fsg beautiful-fsg⟩
'I know that she is beautiful.'

(21) שמעתי שהיא יפה.
shamá-ti ⟨she-hì yaf-á⟩$_{COMP/ISR}$
hear:past-1sg ⟨comp-3fsg beautiful-fsg⟩
'I heard that she is beautiful.' (a general hearsay, not referring to a specific speech act) **or** 'I heard that she was beautiful.' (a specific speech act)

(22) אמרתי לה שהיא יפה.
amár-ti l-a ⟨she-hì yaf-á⟩$_{ISR}$
say:past-1sg dat-3fsg ⟨comp-3fsg beautiful-fsg⟩
'I told her that she was beautiful.'

A speech report can be referred to *en bloc* using the proximal demonstrative *ze* 'this'. Thus, (23) could be a retort to (22):

(23) אתה אמרת את זה ברצינות?
atá amár-ta et **ze** bi-rtsinút?
2msg say:past-2msg acc **prox.msg** in-seriousness
'Did you say that seriously?'

3.2. Reporting verbs

The reporting verb usually appears **before** the speech report, although in literary style, it can follow the speech report either immediately or after the A, i.e., either *'Go away!', **said** the child* or *'Go away!', the child said*, the **former** being of a higher register.

The most common verb used in both DSRs and ISRs is the transitive *amár* 'say:3msg.past'. This verb has suppletive future and infinitive forms: *yagíd* 'say:3msg.fut' and *le-hagíd* 'inf-say' respectively. That said, the future and infinitive forms *yomár* and *l-omár* exist but, unlike in Hebrew, they are not normally used in Israeli.

As previously seen, Israeli has a plethora of other reporting verbs (see Table 1, overleaf).

Moreover, colloquial Israeli often employs *asá*, lit. 'do:3msg.past', as a reporting verb:

(24) אז המהבולה הזאת עושה לי: "צ'מע קטע! אני מה זה בדאון"
אז עשיתי לה: "למה מה קרה? מי מת? בואי לסרט, תום אומר שזה פצצות לגבות!"

az ha-mahabúl-a ha-zòt **osá** l-i
so def-fool-fsg def-prox.fsg **do**:fsgpres dat-1sg

[chmá kèta, anì má ze be-dàwn]$_{DSR}$
[hear:2msgimp fragment 1sg what prox.msg in-down]

az **asíti** l-a [làma má karà? mí met?,
so **do**:1sgpast dat-3fsg why what happen:3msg.past who die:3msg.past

bói l-a-séret, tom omér
come:2fsg.imp to-def-film Tom say:msg.pres

⟨she-zé ptsats-ót l-a-gab-ót⟩$_{ISR}$]$_{DSR}$
⟨comp-proxmsg bomb-fpl to-def-eyebrow-fpl⟩]

'So that idiot (f) goes: "Listen, I'm really down". So I was like: "What the hell? What's your deal? Come to the film, Tom says it's wicked." '

Literally: 'So this idiot (f) **does** to me: "Hear a fragment, I'm what in a down!". So I **did** to her: "Why, what happened? Who died? Come to the film, Tom says that this is bombs to the eyebrows!".'

Table 1

Classification of Reporting Verbs

Verb	Transitivity	Translation	Semantic Class
amár	tr	say	saying
sipér	tr	tell, recount (cf. *safár* 'count')	saying
hodía	tr	announce, notify	saying
yidéa	tr (O=addressee)	inform	saying
taán	tr	claim	saying
hitsír	tr	declare	saying
hikhríz	tr	proclaim	saying
tsaák	tr	shout	saying + *speech manner*
lakhásh	tr	whisper	saying + *speech manner*
milmél	tr	mutter	saying + *speech manner*
shar	amb	sing	saying + *speech manner*
gimgém	intr	stutter	saying + *speech manner*
tsavákh	tr	scream	saying + *speech manner*
tsarákh	tr	yell	saying + *speech manner*
zaák	tr	cry out	saying + *speech manner*
tsahál	intr	rejoice	saying + *speech manner*
yilél	intr	howl	saying + *speech manner*
yibév	intr	wail	saying + *speech manner*
tsikhkék	intr	giggle	saying + *speech manner*
shaág	intr	roar	saying + *speech manner*
teér	tr	describe	proposition
hisbír	tr	explain	proposition
tsién	tr	mention	proposition
hizkír	tr	mention	proposition
heelíl	tr	allege	proposition
heíd	intr	testify	proposition
diveákh	tr	report	report
perét	tr	detail	report
hosíf	tr	add	report
hivtíakh	tr	promise	promise
iyém	intr	threaten	promise
hizhír	tr (O=addressee)	warn	promise

Table 1 (cont.)

Verb	Transitivity	Translation	Semantic Class
makhá al	intr	protest about	complaint
hitlonén	intr	complain	complaint
kitér	intr	whine, whinge (cf. Polish Yiddish *kútər* 'male cat, whiner')	complaint
kavál al	intr	complain about (high register)	complaint
hitonén	intr	complain	complaint
tsivá	tr (O=addressee)	order	command
horá	intr	order	command
pakád	tr (O=addressee)	command	command
amád al kakh	intr	insist	command
hiftsír (be)	intr	urge (high register)	requesting
bikésh (mi)	tr	request (from)	requesting
darásh (mi)	tr	demand (from)	requesting
shaál	tr (O=addressee)	ask	asking
tahá	intr	wonder	asking
aná	intr	answer	answering
heshív	intr	reply	answering
hegív	intr	react	answering
gaár	intr	scold	rebuke
nazáf (be)	intr	reprimand	rebuke
hokhíakh	tr (O=addressee)	reprove	rebuke
odéd	tr (O=addressee)	encourage	encouragement
shamá	tr	hear	hearing

3.3. DSR without a reporting verb

Whereas a reporting verb is obligatory in ISR, it is possible to have a DSR without it:

(25) חוקרי המשטרה לנער הערבי: "למה האיבר מין שלך עטוף בבד לבן?"
הנער לחוקרים: "אמא שלי אמרה לי שהוא צריך להיות מוכן לשבעים ושתיים حور عين".

khokr-éy ha-mishtará l-a-náar ha-arav-í:
investigator-mpl:constr def-police dat-def-teenager def-Arab-msg

[láma ha-evàr mín shel-khà atúf be-bàd laván?]$_{DSR}$
[why def-organ sex gen-2msg wrap:msg.pres:pass in-cloth white]

ha-náar l-a-khokr-ím:
def-teenager dat-def-investigator-mpl

[íma shl-ì amr-á l-ì
[mother gen-1sg say:3past-fsg dat-1sg

⟨she-hù tsaríkh li-yót mukhán
⟨comp-3msg need:msg.pres inf-be ready:msgpres:pass

le-shivím ve-shtáim *ħu:r ʕi:n*⟩ISR]DSR
to-seventy and-two *ħu:r ʕi:n*⟩]

'The police investigators to the Arab teenager: "Why are your genitals wrapped with white cloth?" The teenager to the investigators: "My mother told me that they needed to be ready for the seventy-two *ħu:r ʕi:n*"'[1]

4. Direct speech report as a basis for lexical derivations

Israeli has many NPs which are lexicalized DSRs, usually couched in the first person:

(26) ה"אני מאמין" של מבקר המדינה הנכנס: מלחמה עיקשת בשחיתות.
ha-[**aní maamín**]NP shel mevakér ha-mdiná ha-nikhnás:
def-[**1sg believe:msg.pres**] gen comptroller def-state def-enter:msg.pres

milkhamá ikésh-et b-a-shkhitút
war:fsg stubborn-fsg in-def-corruption
'The incoming State Comptroller's credo: unrelenting war on corruption.'

(27) הח"כ יצא ב"אני מאשים" נגד הממשלה.
ha-khá-k yatsá
def-M(ember)-K(nesset) come.out:3msg.past

be-[**aní maashím**]NP néged ha-m(e)mshalá
in-**1sg** **accuse:msg.pres** against def-government
'The MK (Member of Knesset (Israeli Parliament)) came out with a *J'accuse* against the government.'

[1] Seventy-two حورعين [ħu:r ʕi:n] are promised to the faithful martyrs in Suras 44 : 54 and 52 : 20 of the Koran. Muslims believe that these are "dark, wide-eyed (maidens)", **virgins**. However, Luxenberg (2000) suggests that *ħu:r ʕi:n* are actually "white (grapes), jewels (of crystal)". In other words, Muslim martyrs will not get virgins but **sultanas**(!), the latter with the meaning of white raisins/grapes. Syriac [ħu:r] 'white (fpl)' is associated with 'raisin'—cf. Zuckermann (2006a). If this alternative interpretation is true, or rather, if one can convince fundamentalist Muslims that it is true, it has the potential to change the course of history, at least in cases like the above true story of a Palestinian teenager caught in Israel just before attempting a suicide-bombing.

Consider also the following:

(28) גישת ה"יהיה בסדר" ותופעת ה"סמוך עליי" עוד יהרסו את צה"ל.
gishá-t ha-[**yiyé be-séder**]
attitude-constr def-[**be:3msg.fut in-order**]
ve-tofaá-t ha-[**smókh al-áy**]
and-phenomenon-constr def-[**trust:2msg.imp on-1sg**]
od yaharsú et tsáhal[2]
yet destroy:3pl.fut acc IDF
'The "she'll be right" attitude and the "trust me!" phenomenon may destroy the IDF (Israel Defence Forces).'

5. Functional, stylistic and discourse preferences

DSR is common in Israeli in informal speech or story-telling, and is often employed in jokes; in-your-face Israelis often use (sometimes macabre) self-deprecating humour:

(29) האמריקאי אומר לאשתו: "בבקשה תעבירי את הסוכר, סוכריה!"
הבריטי אומר לאשתו: "בבקשה תעבירי את הדבש, דובשנית!"
הישראלי אומר לאשתו: "תעבירי את הסטייק, יא פרה!"
ha-amerikáy omér le-isht-ó
def-American:msg say:msg.pres dat-wife-3msg.poss
[bevakashá taavír-i et ha-sukár, sukaryá!]$_{DSR}$
[please pass:2fut/imp-fsg acc def-sugar candy]
ha-bríti omér le-isht-ó
def-Briton:msg say:msg.pres dat-wife-3msg.poss
[bevakashá taavír-i et ha-dvash, duvshanít!]$_{DSR}$
[please pass:2fut/imp-fsg acc def-honey honey.cookie]
ha-israelí omér le-isht-ó
def-Israeli:msg say:msg.pres dat-wife-3msg.poss
[taavír-i t-a-stek, ya pará!]$_{DSR}$
[pass:2fut/imp-fsg acc-def-steak voc:derog cow]
'The American tells his wife: "Would you pass the sugar, sugar!". The Briton tells his wife: "Would you pass the honey, honey!". The Israeli tells his wife: "Pass the steak, *ya* cow!" '.

[2] Acronym of ***tsvá ha-haganá le-israél*** 'Israel Defence Forces'.

The following passage employs both DSR and ISR. DSR is employed for immediacy and engaging effect, including a couple of serendipitous *ex postfacto* puns. ISR conveys important new information but since the specific form of the speech act is not crucial, it is backgrounded:

(30)

הזונה המולדבית שאלה את הצעיר הפלסטיני: "בא לך להתפרק?"
הוא ענה לה: "אין לי ראש לזה!"
רק אחר כך השוטרים סיפרו לה שהוא בעצם היה המחבל המתאבד.

ha-zoná ha-moldávi-t shaal-á et ha-tsaír ha-falestín-i:
def-prostitute def-Moldavian-fsg ask:past-3fsg acc def-youth def-Palestinian-msg

[bá le-khà le-hitparék?]$_{DSR}$
[come:msgpres dat-2msg inf-disassemble]

hu aná l-a: [én l-i rósh le-zè!]$_{DSR}$
3msg reply:3msg.past dat-fsg [exis.cop:neg dat-1sg head dat-prox.msg]

rak akhár kakh ha-shotr-ím siprú l-a
only after so def-policeman-mpl tell:3mpl.past dat-fsg

⟨she-hù be-étsem hayá ha-mekhabél ha-mitabéd⟩$_{ISR}$
⟨comp-3msg in-substance be:3msg past def-terrorist def-suicider⟩

'The Moldavian prostitute asked the Palestinian youth: "Do you feel like getting off?" (lit. "Does it come to you to disassemble?"). He replied: "I'm not in the mood!" (lit. "I don't have a head for it!"). Only later did the policemen tell her that he actually was the suicide bomber.'

6. Concluding remarks

There is a clear distinction in Israeli between direct and indirect speech. The indirect speech report, which is a subset of complement clauses, is characterized by a shift in person, spatial and temporal deixis. However, unlike in English, the verbs usually do not undergo a tense shift. Israeli has various lexicalized direct speech reports. By and large, Israeli reported speech constructions reflect Yiddish and Standard Average European patterns, often enhancing a suitable pre-existent Hebrew construction.

Abbreviations

1 = 1st person; 2 = 2nd person; 3 = 3rd person; A = transitive subject; acc = accusative; amb = ambitransitive; CAUS/caus = causal; CC = copula complement;

comp = complement(izer); constr = construct-state; COP/cop = copula; CS = copula subject; dat = dative; def = definite; derog = derogatory; DSR = direct speech report; E = extended intransitive; exis = existential; f = feminine; fut = future; gen = genitive; imp = imperative; INF/inf = infinitive; INTER = interrogative; intr = intransitive; ISR = indirect speech report; m = masculine; neg = negator/negative; NP = noun phrase; O = transitive object; pass = passive; pl = plural; poss = possessive; pres = present; prox = proximal demonstrative; S = intransitive subject; sg = singular; tr = transitive; V = verb; VCC = verbless clause complement; VCS = verbless clause subject; voc = vocative.

References

Luxenberg, Christoph 2000. Die syro-aramäische Lesart des Koran. Ein Beitrag zur Entschlüsselung der Koransprache. Das Arabische Buch, Berlin.

Zuckermann, Ghil'ad 1999. Review article of Nakdimon Shabbethay Doniach and Ahuvia Kahane (eds), The Oxford English-Hebrew Dictionary. In: International Journal of Lexicography 12:325–46.

Zuckermann, Ghil'ad 2003. Language contact and lexical enrichment in Israeli Hebrew. Palgrave Macmillan, London.

Zuckermann, Ghil'ad 2006a. haivrít kemítos [Hebrew as myth]. Am Oved, Tel Aviv.

Zuckermann, Ghil'ad 2006b. A new vision for Israeli Hebrew: Theoretical and practical implications of analysing Israel's main language as a semi-engineered Semito-European hybrid language. In: Journal of Modern Jewish Studies 5:57–71.

INFORMATION FOR CONTRIBUTORS

Please provide typewritten double-spaced manuscripts with wide margins (50 characters per line, 25 lines per page). Type on one side of the paper only and number all the pages consecutively. The following information should be provided on the first page: the title of the article (or, in case of reviews, the title of the book under discussion, the name of the publisher, the place and date of publication as well as the number of pages), the author's full name and address (including email address) and an abbreviated title of the paper (not exceeding 45 characters including spaces). Only original papers will be published and a copy of the *Publishing Agreement* will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement. Authors are requested to send two hard copies of their manuscript + a floppy disk under DOS/Windows. The manuscript should be accompanied by a list of five keywords and an abstract of about 100 words. It should be typed on a separate sheet of paper. Tables, diagrams and illustrations (with the author's name and an Arabic number) should be presented on separate sheets. Captions should be typed on a separate sheet and placed at the end of the manuscript. Footnotes should be typed double-spaced on a separate sheet at the end of the article. All language examples (and only these) are to be italicized (single underlining). Citations in the text should be enclosed in double quotation marks (" " in case of a paper written in English, „ " in German and « » in French).

Reference to a publication should be made by the name of the author, the year of publication and, when necessary, by page numbers in the following ways:

...as described by Schmidt (1967)...
...as referred to by Hoover (1967, 56–78; 1976, 43)...
...mentioned by several authors (Elgar 1978, 77; Williams 1981, 154–6).

An alphabetically arranged list of references should be presented at the end of the article as follows:

Bárczi, Géza 1958a. Magyar hangtörténet [The historical phonology of Hungarian]. Akadémiai Kiadó, Budapest.

Bárczi, Géza 1958b. A szótövek [Word stems]. Akadémiai Kiadó, Budapest.

Brockhaus, Wiebke 1995. Skeletal and suprasegmental structure within Government Phonology. In: Jaques Durand – Francis Katamba (eds): Frontiers in phonology: Atoms, structures, derivations. 180–221. Longman, Harlow.

Cole, Jennifer 1995. The cycle in phonology. In: Goldsmith (1995: 206–44).

Goldsmith, John A. (ed.) 1995. The handbook of phonological theory. Blackwell, Cambridge MA & Oxford.

Kaye, Jonathan – Jean Lowenstamm – Jean-Roger Vergnaud 1990. Constituent structure and government in phonology. In: Phonology 7: 301–30.

Tomioka, Satoshi 1997. Focusing effects and NP-interpretation in VP-ellipsis. Doctoral dissertation, University of Massachusetts, Amherst.

For marking subsections decimal notation should be applied. Do not use more than four digits if possible. Examples within the text should be marked in italics. Meanings are to be rendered between inverted commas (' '). If glosses are given morpheme by morpheme, the initial letter of the gloss should be placed exactly below that of the example. Grammatical morphemes can be abbreviated in small case letters connected to the stem or the other morphemes by a hyphen. No period should be applied in such abbreviation. For example:

(1) (a) A sólymaid elszálltak
the falcon-gen-pl-2sg away-flew-3pl
'Your falcons have flown away.'

Examples can be referred to in the text as (1a), (1a–d), etc.

One proof will be sent to the author. Please read it carefully and return it by air mail to the editor within one week.

New subject collections available

Akadémiai Kiadó offers **a new suite of electronic-only journals** collections in Arts & Humanities, Life Sciences, Sciences, and Social Sciences with significant pricing discounts.

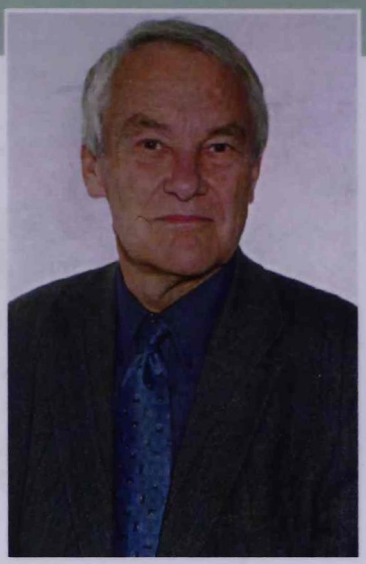
2006

Akadémiai Journals Collection ■ **Arts & Humanities**

Across Languages and Cultures
Acta Historiae Artium *(Online version includes abstracts only)*
Acta Linguistica Hungarica
Acta Orientalia Academiae Scientiarum Hungaricae
Hungarian Studies
Studia Slavica Academiae Scientiarum Hungaricae
Studia Musicologica Academiae Scientiarum Hungaricae
Verbum – Analecta Neolatina

For further details please visit our homepage at
www.akademiaikiado.hu/collections

1828
AKADÉMIAI KIADÓ



FERENC KIEFER, editor-in-chief, professor emeritus
Member of the Hungarian Academy of Sciences
President of CIPL
Corresponding member of the Austrian Academy of Sciences
Member of Academia Europaea and of
European Academy of Arts, Sciences and Humanities
Main research areas:
morphology, lexical semantics, semantics-pragmatics interface
He has published 11 books (in Hungarian, English, French and German) and more than 200 articles

CODEN: ACTA LINGUIST. HUNG.
SIGILLUM: ALHUE 8

ISSN 1216-8076
9 771216 807004

phonology ■ morphology ■ syntax ■ semantics and pragmatics ■ sociolinguistics ■ cognitive linguistics